FUZZY SYSTEMS THEORY AND MEDICAL DECISION MAKING

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

Robert Craig Elder

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Georgia Institute of Technology

December, 1976

FUZZY SYSTEMS THEORY AND MEDICAL DECISION MAKING

Approved:

Augustine O. Esogbue, Ph.D., Chairman

Stuart J. Deutsch, Ph.D.

Walter L. Bloom, M. D.

Date approved by Chairman: December 10, 1976

ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere appreciation to those persons who have aided me in the research and preparation of this thesis.

In particular, I am extremely grateful to Dr. Augustine O. Esogbue, my thesis advisor, for his initial motivation and continued assistance and availability throughout the various phases of this research; to Dr. Stuart J. Deutsch for his interest, opinions, and helpful evaluations of the proposed mathematical modeling; and to Dr. Walter L. Bloom for guiding and encouraging my efforts to master the medical knowledge needed to pursue research in this field.

I would also like to thank the physicians in the Atlanta area who answered questionnaires and granted interviews so that their decision processes might be better studied and analyzed.

Finally, to my family and friends, I want to extend a special thanks for their needed support and understanding. Above all, I express my deepest appreciation to my girlfriend, Donna, for her endless tolerance, patience, sensitivity, and moral support in the past few months. Without her presence and typing abilities, completion of this thesis in the desired time frame would have been impossible.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

This research centers on the use of fuzzy systems theory to more realistically model physician decision making processes. Various physician decision processes, including the general diagnostic and treatment processes, are described and detailed so that their fuzzy aspects might be discerned. Deficiencies of past mathematical representations of both fuzzy and nonfuzzy decision aspects are outlined and evaluated critically.

From these analyzed decision processes, the diagnosis decision process was selected as the most appropriate for mathematical modelling via fuzzy set theory. The diagnosis decision process includes decisions upon medical hypotheses, preliminary diagnoses, and final diagnoses. Patient information including past history, present symptoms, observed signs, or test results are needed to reach these various diagnosis decisions. Fuzzy set theory is primarily used to evaluate this patient information so that it can be incorporated into the proposed diagnosis decision model. The model of the actual diagnosis decision uses cluster analysis techniques to determine the proper diagnosis. The patient's sickness is clustered with the most similar disease stage using a revised Euclidean distance similarity measure.

An illustrative example of the diagnosis decision model is offered to clarify and illustrate fuzzy and nonfuzzy aspects of the proposed model. In this instructive example, the model reaches medical hypotheses for rheumatic valvular heart diseases. The diagnosis

decision model is analyzed and validated through this example. Medical hypotheses reached by a physician and the fuzzy computerized model are contrasted so that problem areas or discrepancies are uncovered. Suggestions are proposed to remedy these discrepancies in future research. As a diagnostic tool, the initial accuracy of the fuzzy diagnosis model was not extremely high, but the results were encouraging enough to warrant further model refinement and applications. The decision model did succeed in representing more of the influential aspects important to actual physician decisions.

CHAPTER I

INTRODUCTION

In recent years, more and more attention has been focused on various problems in the health systems field. Due to increased public demands, government as well as private research organizations have increased their spending in an effort to improve health care and to better understand the problems confronted by people involved in the health fields. Various techniques, including those of mathematical modelling, have been used by physicians, engineers, and mathematicians to analyze different problem areas to obtain these improvements.

One specific problem area, that of dicision making by physicians, has received much attention in recent years. Decisions made by physicians and psychotherapists during the initial interview with the patient, the diagnostic process, and the treatment process have been modelled extensively. A variety of modelling techniques have been used, including those of operations research, decision theory, and pattern recognition.

The physicians' patient oriented decisions are the primary points of concern in this research. The main thrust of this thesis is directed toward the analysis of decisions made by the physicians who act as clinicians, rather than physicians in general. A clinician is the physician in direct contact with the patient, the one who accepts responsibility for helping the patient, and the one who offers

relief to the patient. This eliminates physicians in laboratories, pathologists and others who do not directly influence the patient's care.

Several benefits should emerge from the analyzation and modelling of these physicians' decisions. From an educational viewpoint, physicians receive only limited formal training in the mathematical analysis of decision making. Many believe their decision making is unsusceptible or too complex for such investigation and modelling. As a result, physicians benefit little from the use of mathematics in evaluating the structure, logic and information used in their decisions. If precise, realistic models are created and scientifically presented, physicians will hopefully interpose the benefits into specific applications of decision making. By employing these models in specialized areas of medicine, physicians should strengthen their ability to analyze decisions, recognize contributory factors and evaluate possible alternatives. Due to increasing requirements upon physicians, such mathematical models have often been aimed at directly assisting or eliminating the physician from numerous decisions. This has become more feasible with the advent of computers and an embryonic effort by medical personnel to collect and compile the vast quantities of information needed for data bases. Models directed toward computerized diagnosis, treatment selection, and test evaluation have multiplied and matured over the past decade. Regretfully, these models have notable shortcomings, many of which will be mentioned later in the research.

Purpose of the Research

The objective of this research is to invoke the concepts and methods of fuzzy decision theory in an effort to more realistically model the physicians' decision processes. To attain this, the general diagnostic process, the general treatment process and constraints by the environment are first developed and analyzed in detail. The fuzzy aspects, not usually incorporated in previous models, are specifically identified and exemplified. A summary of previous modelling techniques is given and their deficiencies as well as misrepresentations are outlined.

Based on this analysis, a more realistic model of the diagnosis decision process is presented through the use of fuzzy set theory. Attempts are made to quantify those aspects which are fuzzy, difficult to quantify, or imprecise. For a specialized area of medicine, an illustrative example of this general model is given. Model computerization and validation is then exhibited for diagnoses in this medical field.

Background

Models of physicians' decision processes presented to date use many techniques which are prevalent in classical decision theory. Variations of Bayesian statistics, decision trees, pattern classification techniques and others have been used in modelling these processes. Unfortunately, many fail to include important aspects of the actual decisions. When applied to specific disease areas, the accuracy of such models does not even approach that required of

physicians [23]. Thus, more detailed analysis of these decision processes and more precise, realistic models are needed to produce accurate results for educational or applied purposes.

To accomplish this, attempts are made to alleviate certain discrepancies of past modelling techniques. One severe shortcoming of these past techniques is their inability to differentiate between fuzziness and randomness in the physicians' decision processes. Randomness deals with the uncertainty of membership in a well defined set, and can be effectively represented using probability theory. But many aspects of these decision processes are intrinsically fuzzy, not random. If the boundary of the set cannot clearly be defined, then certain aspects of the decision cannot be represented as being included in the set or excluded from the set. Thus, the aspect becomes a member of a fuzzy set and is represented by a degree of membership in that set. Clearly, techniques being used to represent randomness should not be used to represent this fuzziness. Unfortunately, in many instances, past models have done just that. To remedy this situation, fuzzy aspects are illustrated throughout the analysis of the diagnostic and treatment process. Fuzzy set theory is then used to model such aspects in the diagnosis decision process.

Study of Literature

The literature relevant to this research on physician decision making has been restricted to the following main areas:

1. Medical references.
2. References on mathematical modelling, applied to

physician decision making.

3. References on fuzzy set theory.

These literature groups comprise the formal information used in the analysis, modelling, and simulation of the diagnostic and treatment processes.

Medically oriented references on physician decision making and cardiovascular diseases have been used extensively in the systems description phase as well as in the example of the resulting mathematical model. Literature oriented towards the decision processes specifically involves the initial interview with the patient, the diagnostic reasoning process, and the test selection process. In his work concerning the patient's initial interview, Stevenson [15] details in depth the interactive events between physician and patient. Relevant aspects of the patient's history, such as previous illnesses, past medical disorders, and family background, are thoroughly discussed. Difficulties encountered by the physician during the interview, variations and contrasts in interviews, and the physician's ability to guide and control the interview also receive specific analytical attention.

Research by Krieg, Gambino, and Galen [10] gives insight into the proper medical basis for diagnostic test selection. Measures of test usefulness, outlined later in this research, are proposed for determining the medical effectiveness of a particular test. Due to ever increasing costs, the need for a test is also questioned whenever the results may be nonessential to the diagnosis decision being made. Informal criteria for evaluation of this need is therefore proposed.

The most thorough medical analysis of the diagnostic and

treatment decision processes is offered by Feinstein. His primary work [3] illustrates in depth such aspects of clinical judgment as symptom interpretation and designation, sign observation, taxonomy problems encountered when reaching a diagnosis, and treatment and therapy considerations and selection. His other related research [4,5] gives one of the few detailed insights into the diagnostic reasoning process used by physicians to explain signs, symptoms and other attributes when converging on a diagnosis. Criticism of past mathematical techniques used to represent physician decision making prevails in his works, since oversimplification, omission or misrepresentation by researchers alters the actual decision processes.

The second major group of medical references contains information used for the illustrative example of the diagnostic model. This information centers on heart disease, but more specifically rheumatic valvular disease. The literature from this category is basically similar in content and is distinguishable only by the quantity and quality of information relevant to the example. The most useful research in this specialized field is attributable to Reichek, Shelburne, and Perloff [13]. Each of the rheumatic valvular diseases is detailed with respect to cause, disease development, and treatment. The patient's past history and present symptoms, needed for the model example, are enumerated and interrelated. Similar works have also been presented by Conn and Horowitz [2] and Hurst [8]. Both of these are somewhat voluminous, since all cardiovascular diseases and disorders are discussed, but the information conveyed remains pertinent.

Mathematical modelling of physicians' decision processes

encompasses the next group of references. Comprehensive research by Croft [23] cites the most prevalent computerized diagnosis techniques available at the present time. Assumptions these models require and problems confronted by their use for computerized diagnosis are explicated. Variations on Bayesian and clustering models are listed and explained. These approaches are then evaluated with respect to their accuracy in reaching the proper diagnosis.

Jacquez [34] has assembled an excellent collection of modelling theory and applications in the diagnostic decision domain. This volume is essentially the Proceedings of the Second Conference on the Diagnostic Process held at the University of Michigan in 1971. Empirical and statistical diagnostic models are therein elucidated and possible heuristic approaches are outlined.

A somewhat new and creative approach to the analysis of the initial physician decisions has been introduced by Bellman and Smith [19]. Although the analysis and model deal with the initial psychotherapeutic interview, the mathematical approach is still very relevant to initial interviews by other physicians as well. The systems analysis of this adaptive decision process defines possible systems, influences and responses which are pertinent to the interview and its results. The resultant simulations found in the final phases of this work are very similar to patient-physician interactions of history and symptoms modelled later in this research.

Chesler, Hershdorfer and Lincoln [22] have developed a systems analysis of the diagnostic and treatment processes, aimed at the use of specific clinical information. Their results are somewhat similar

to those uncovered in the systems description in Chapter II, even though research for this thesis was performed without knowledge of the above reference. The information used for their research was gained without extensive physician input and lacks much of the medically oriented data found herein.

There exists an abundance of mathematical applications for decisions made during different phases of the diagnostic and treatment processes, many of which are similar to those mentioned above. Since most of these will be noted and evaluated in Chapter II of this research, additional comments on previous models will be delayed to this time.

The last section of literature, relevant to the modelling of the initial phase of the diagnostic process, concerns fuzzy set and fuzzy systems theory. The concept of a fuzzy set was first introduced by Zadeh [60] in 1965 to differentiate between randomness and fuzziness. Since that time, references on fuzzy sets have increased dramatically as more and more theory and applications have been developed. An early contribution due to Bellman, Kalaba, and Zadeh [56] develops a framework for abstraction and pattern classification through fuzzy sets. Based on sample data of a fuzzy set, a general membership function can be constructed to represent the fuzzy set as a whole. Then membership of a new element can be classified according to this general membership function.

Kochen [58,59] and Zadeh [63] have introduced and developed the application of fuzzy set theory to fuzzy semantics, fuzzy adjectives, and psychology. This was done via prominent periodicals and recently at the U.S. Japanese Seminar on Fuzzy Sets and Their Applications.

Techniques used in these references are somewhat similar to those incorporated in this research to quantify fuzzy symptom descriptions in the initial information gathering session. Included in these references are experimental results verifying the true fuzziness of specific, imprecise terms.

Another application of fuzzy theory, initiated by Bellman and Zadeh [52], has been in the area of decision making in a fuzzy environment. For this problem, the system under consideration is nonfuzzy (real numbers), while constraints and goals are fuzzy in nature, requiring the intersection of fuzzy constraint and goal sets to obtain the solution set. Although the system under consideration in this research is also fuzzy, certain notions discussed by Bellman and Zadeh seem applicable to this effort. Fuzzy allocation processes, investigated and developed by Esogbue [55,56], give further insight into the various uses of fuzzy theory. Examples are given for the evaluation of membership functions, fuzzy constraints, and fuzzy objectives. Applications of this model to the allocation of funds for cancer research have been suggested and deemed feasible.

The development of the theory of fuzzy sets has grown rather rapidly since its conception. Included in these developments is a variety of theory, the most directly relevant of which deals with such concepts as the entropy of a fuzzy set [54], shadow of a fuzzy set [61], similarity relations and fuzzy ordering [64], and L-Fuzzy sets [57]. Certain specific aspects of fuzzy theory are given as an appendix to this research.

CHAPTER II

MEDICAL DECISION MAKING: A SYSTEMS DESCRIPTION

In the following sections, the physicians' decision processes are analyzed in detail. The general diagnostic process, the general treatment process, and limitations imposed by the environment are extensively discussed. Accompanying each of these discussions is a brief summary and evaluation of past mathematical techniques used to model these decision processes. The information needed for this descriptive analysis was obtained from key medical journals and references, followed by mail questionnaires and personal interviews with local physicians.

The General Diagnostic Process

The diagnostic decision process is a sequence of decisions made by a physician in at attempt to identify and explain the ailments, disorders, and diseases present in a particular sick patient. This process involves the acceptance of the patient into the physician's care, and the collection and evaluation of pertinent information at various intermittent stages. Such information, obtained through discussion, observation, and tests, is significant to the convergence upon effective preliminary and final diagnoses. To study the general diagnostic process, this analysis has been structured into four subsystems, which are illustrated in Figure 1. The interactions between these subsystems are also shown in the same diagram. These interactions, as well as the

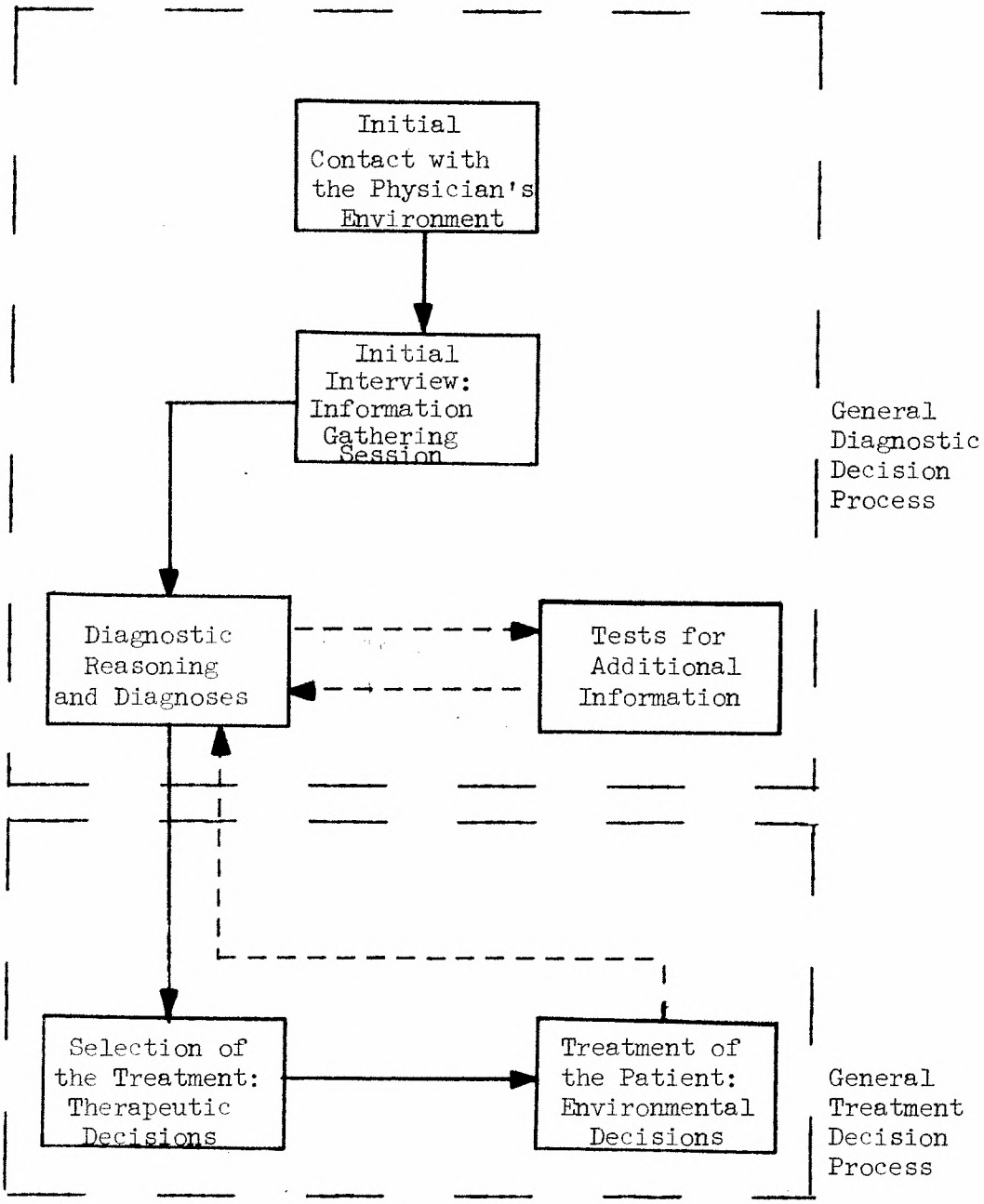

Figure 1. Systems Diagram of the General Diagnostic and Treatment Processes.

subsystems themselves, are discussed in individual sections later in the chapter.

In an effort to clarify the meaning of certain terms used in this portion of the research, the following definitions are given:

Attributes or Manifestations - signs, symptoms, pieces of pertinent history, test results, or other bits of information known about the sick patient.

Symptoms - the patient's subjective feelings or discomforts associated with his present sickness.

Signs - the findings observed by a physician during physical examination of the sick patient.

Test - any clinical activity used for obtaining additional information about the sick patient (E.C.G., X-Rays, tests on body fluids and tissue, and other methods evaluating the state of the patient).

Diagnosis - the identifying of a disease, group of diseases, or other reason that explains the occurrence of the patient's attributes.

Hypothesis - a medical conjecture regarding the disease or cause of disease in the patient, based solely upon the past history and present symptoms.

Preliminary Diagnosis - an early diagnosis, initially made following physical examination of the patient and modified by clinical and diagnostic test results.

Final Diagnosis - the endpoint of the diagnostic process, upon which treatment or therapy is based.

Initial Contact With the Physician's Environment

This initial subsystem of the diagnostic process involves the sick patient contacting the physician's environment (home, office, hospital, etc.) in an effort to obtain some sort of medical assistance. This subsystem includes the acceptance of a new patient into the physician's care, as well as the renewed care of a previous patient with an undiagnosed sickness. The sick patient contacts the physician's environment either by entering his office or by telephoning his home, office or hospital.

With the exception of the emergency room of a hospital, relatively few patients enter the physician's office without an appointment of some sort. According to the results of the questionnaire, less than ten percent of the patients seen by each physician entered his office without a previous appointment. Some physicians accept patients only by appointment, so the patient may be turned away from the physician's office altogether. If a patient enters a physician's office seeking assistance, he may be turned away or referred to another physician, he may be given a future appointment, or he may wait and see the physician directly. Whenever the patient comes into direct physical contact with the physician, he enters into our next decision subsystem.

The majority of sick patients contact the physician's office, hospital, or home over the telephone. If the patient has been accepted into the physician's care, he may desire only a future appointment, or may want or need more immediate relief.

Given that this patient seeks immediate relief over the telephone and his sickness has not been observed or diagnosed by a physician,

a special case of the diagnostic and treatment process emerges. First of all, the physician must make his decisions based solely on obtainable information. In this case, the obtainable information is limited to the patient's past history and the symptoms related over the telephone. Thus, the physician develops a hypothesis, preliminary diagnosis, or final diagnosis based primarily on verbal interactions between physician and patient. Note that this verbal interaction will be discussed in a subsequent subsystem.

From comments in the questionnaire, most physicians in this situation would grant an appointment time reflecting the urgency associated with the hypothesis, preliminary diagnosis, or final diagnosis. If the physician determines emergency relief or hospitalization is necessary, the patient is either met by the physician at the hospital or the physician recommends emergency room treatment. If no emergency exists, the earliest appointment time would depend upon the needs of the patient and the availability of the physician.

For these telephone cases where the physician is unable to see the patient soon, and no emergency exists, certain relief decisions must be made. The results of the questionnaire revealed that two-thirds of the physicians would offer prescription drugs as relief if the symptoms warranted. Prescription drugs mentioned by specific physicians included antibiotics, antihistamines, analgesics, and pain medication. Factors influencing the physician's selection of prescription drugs include familiarity and knowledge of the patient, pain and discomfort of the patient, and the physician's interpretation of the urgency of the situation. Because of lack of information or lack or urgency, the physician

may offer no immediate prescription relief and request to see the patient during his office hours. Approximately one-half the physicians questioned offered nonprescription relief to their patients warranting it. This relief seem to be temporary in nature and examples include antipyretics (aspirin), laxatives, and cough medicine. Three out of five physicians offering prescription or nonprescription relief in this situation occasionally did not recommend to see the patient at a later date. Factors given that influence this decision include knowledge of the patient, lack of danger or severity, and physician's satisfaction with the diagnosis and treatment selection. Thus, in these instances, diagnostic decisions have been made without physician observation or test results, and the treatment selection was deemed satisfactory and permanent.

Since this subsystem introduces the general diagnostic decision process, most of the decisions are procedural and lack the structure or need for modelling. The preceding special case was analyzed in this section because the patient never came into direct contact with the physician. Modelling techniques used for the general diagnostic and treatment processes, discussed in subsequent subsystems, can be modified to deal with this special case. Evaluation of these modelling techniques will also be given at that time.

Initial Interview: Information Gathering Session

This stage of the diagnostic process begins with the patient's initial direct contact with the physician. During this subsystem the physician attempts to gain information about the state of the sick patient, on which he can base his diagnosis. The following information can be obtained in this initial interview:

1. The patient's history
   a. Past history
   b. Symptoms of the present sickness related to the physician by the patient.
2. The signs observed by the physician while performing the physical examination.

The past history supplies data about the specific patient or host of the disease. The signs and the symptoms are interpreted as interactions between the host or patient and the disease. The only other information available concerns data on the disease itself and is obtained in the latter subsystem of clinical and diagnostic testing.

The quality and quantity of this pertinent information obtained during the initial interview influences the possibility of a correct or incorrect diagnosis. During this interactive process between the patient and the physician, it is important that the information be as precise as possible, given the constraints of the subsystem. One major constraint, imposed by the physician, involves the limitations of time. The questionnaire revealed that almost every physician had an appointment time of fifteen minutes in length, with variations allowed for special examinations. Thus, the lack or availability of this time plays an important role in determining the number and preciseness of the patient's attributes. Other such constraints of this subsystem are given in detail when they are encountered.

The first information relevant to this subsystem is the patient's past history. Each physician questioned during this research required a past medical history from his new patients. Their answers also

revealed that this history taking is usually performed by themselves, rather than by their office personnel. Information requested for the medical history usually includes name, age, sex, religion, racial or national origin, occupation, residency, education, marital status, family background, physical characteristics, past health, and previous diseases and disorders. This history gives a good basis for future medical care as well as insight into the past state of health of the patient. Frequent illnesses, similar sicknesses, and past diseases often relate to present and future medical problems. Occupation, national origin, education, residency, and family background directly influence physical and emotional factors relevant to the well being of the patient. The patient's reactions and answers to pertinent questions also gives a physician a better understanding of the patient and his communicative abilities. Psychotherapists and psychiatrists use much of this information to a greater extent than other physicians in their analysis of the patient.

If the patient has received assistance and treatment before, his medical history is already known by the physician. When the physician meets with the patient, he notes this past history and uses this information for decisions in the present diagnostic process. Past symptoms, signs, test results, treatment, and disorders may be present or important to the current sickness. Other information the physician has at this time includes relief he may have offered the patient for this sickness over the telephone, as detailed in the previous subsystem. Also, many physicians record some information about the patient's sickness or needs at the time the patient makes an appointment, giving

them some idea of the possible problems and decisions that might arise.

After the patient's past history has been taken or noted, the symptoms of the sickness related to the physician by the patient. This relation is a dynamic interaction since the number and preciseness of these symptoms develop or increase as the interview progresses. Often the patient is not with the physician when he is experiencing these symptoms, so the symptom description becomes extremely important. Usually patients relate the most discomforting symptoms first, so the physician must properly question the patient to insure a complete list of symptoms. Physicians admit that patients often describe these symptoms in imprecise, unclear or fuzzy terms, but through verbal and nonverbal physician-patient interaction this impreciseness will hopefully decrease. When detail is lacking, distinguishing features of symptoms are vague and discrete disorders become blurred and confused with each other. Thus, these imprecise symptoms should be clarified into nonfuzzy terms by the patient, if possible. But often the best description of a symptom still remains vague and imprecise. Occasionally, a physician may ask a patient to perceive a symptom over a period of time to more clearly describe the discomfort as well as its location, frequency, or degree of discomfort.

The questionnaire identified several factors influencing the patient's ability to relate symptoms in a precise and clear way. An attempt was made to have these factors quantified according to their influence, but the physicians too often disagreed as to the importance of each of these factors. The variation of each factor in the rating scheme was such that the only precise conclusions possible would be

that the factors varied in importance from physician to physician or that the factors were not susceptible to such a grading by physicians. These factors can be classified into the following three categories:

1. The patient's lack of knowledge regarding medical terminology used to precisely describe symptoms.
2. Factors influencing the patient's ability to communicate.
3. Constraints and limitations of the physician.

The first of these categories involves the patient's lack of familiarity with medical terms used to describe symptoms. The limitations of everyday language and absence of common nomenclature between physician and patient hinders the definition and interpretation of symptoms. The second category deals with personal characteristics of the patients such as education, socio-economic background, shyness, modesty, apathy or indifference, unfamiliarity with the surroundings, and other patient factors affecting symptom description. The most influencial of these seemed to be the patient's education, socio-economic background, and age. The last category includes constraints imposed by the physician or limitations of the physician himself. Constraints imposed by physicians include the time constraint mentioned earlier, as well as physician thoroughness. Physician limitations include the physician's ability to relate to the patient, ask the proper questions, and seek honest and complete answers.

Aside from a description of the symptom itself, several other aspects of the symptom are related in this physician-patient interaction. These aspects make up what is called the <u>symptom profile</u>. Included in this symptom profile are the following aspects of the

symptom: degree of severity, location, longevity, and continuity.

The severity of symptoms frequently varies. Physicians noted that this variation may influence or change the final diagnosis or treatment of a patient. For instance, pain in a particular location may vary over a wide spectrum to change the symptom's identification with a particular ailment. Many symptoms fluxuate, often increasing in severity, reach a peak of maximum severity and then decline. The degree of severity for a symptom of this type could give insight into the disease development in individual patients.

Patients usually describe this degree of severity using terms such as bad, severe, mild, sharp, and other descriptive adjectives. These terms seem rather fuzzy in nature and lack any obvious quantitative structure. But in reaching a diagnosis, these levels of severity must be defined and interpreted by the physician. Note that in some instances the description and degree of severity of a symptom both lack preciseness and clear definition.

The next aspect of the symptom profile concerns the location of the symptom. Generally, location is the most important detail of a symptom since it can determine whether a symptom exists as non-contributory to critical. Some symptoms, such as fatigue, weakness, and depression cannot be associated with any specific location of the body, while others, such as chest pain, may only be limited to a generalized region.

The last two factors of the profile are symptom longevity and continuity. The longevity of the symptom is defined as the length of time the patient has experienced or observed the given symptom. This

aspect may be an indicator of the progress and severity of the disorder or disease in the patient. The continuity factor determines whether the symptom is continuous or intermittent in occurrence. If the symptom occurs intermittently, then the frequency, duration, sudden or gradual onset and sudden or gradual recovery should be noted. Other factors, such as sleep, fatigue, and eating, that accompany the incidence of the symptom can also be distinguished.

The physician must interpret and determine these specifics which make up the symptom profile. His careful questioning should eliminate much of the fuzziness normally associated with patients' descriptions. At the end of the symptom explication, the physician designates a medical name accounting for each sensation and the specifications (profile) related during the interview. This medical designation is usually nonfuzzy in nature. However, the severity quantification terms remain vague and imprecise, similar to the descriptive adjectives used by patients. After the physician finishes designation of the symptoms and profiles, he forms a hypothesis on which he bases his physical examination.

The last information gained from the initial interview comes from the physician's observation of the patient's signs. This includes visual and physical examination of the patient, using basic physical examining instruments, such as stethoscopes, thermometers, and others, to aid in identifying, medically designating, and qualifying these signs. Some symptoms become signs when the physician can observe what the patient has verbally explicated. Any information gained by the patient description can be added to the physician's own perceptions.

Most physicians place more weight on the signs they observe, rather than relying on patients' disclosure of symptoms.

Based on his medical experience and education, the physician classifies his observations as being normal or abnormal. A symptom is usually considered somewhat abnormal since the patient perceives it as being different from what is normal for him. But the physician must determine if the sign he observes is normal for the patient based on his limited knowledge of the patient, as well as his past experience as to what is normal. Making this situation more complex is the fact that signs act very similar to symptoms when it comes to variations in severity. For signs such as burns and congestion, the variation of severity can be enough to affect the final diagnosis or treatment. This makes it essential for the physician to determine an often fuzzy or unclear degree of abnormality for the observed sign.

A sign profile similar to a symptom profile must also be created to give a more detailed description of the specific sign. The important detail of this profile, besides the degree of severity mentioned above, is the location of the sign. From his observations, the physician can usually limit the location to a fairly precise domain by either seeing, feeling, or hearing the sign. Other factors such as longevity and continuity are also included in this profile. Unfortunately, these aspects have to be disclosed by the patient, unless the physician can estimate them based on his observation of the sign's progress in development.

Whereas the preciseness of symptoms is affected by proper communication and other factors, the specification of signs depend on

proper observation and designation. When questioned, physicians admitted that they had incorrectly identified or missed observing many signs. Reasons for this included the following:

1. The observed signs did not warrant further investigation.
2. The physician was inexperienced at the time.
3. The sign rarely occurred in normal practice.
4. Observations were restricted to a limited area.
5. The sign was not significant at the time of observation, but became more profound as time progressed.

Other, more general reasons included error in judgment, misrepresentation, and incorrect evaluation of the observation.

During the development of symptom and sign profiles, the physician must evaluate each of these attributes, eliminating those which do not seem pertinent to the present sickness. Although symptoms and signs are abnormal, they may exist in the host patient continually and not be associated with the present sickness. Others may be of such a nature that they cannot be related to other attributes or any disease. These unimportant symptoms and signs should be eliminated to discourage wasted additional effort and consideration.

At the end of this subsystem, the physician must reach an initial preliminary diagnosis reflecting the patient's pertinent past history, symptoms, and signs. This decision guides the diagnostic reasoning process and test selection process, discussed in the following subsystems.

The mathematical representations of this subsystem have been

somewhat incomplete and oversimplified in past research. First, very little attention has been focused on the physician's proper interpretation of a patient's description of symptoms. Most models begin with medically designated signs and sumptoms, which have already been evaluated by a physician. More mathematical emphasis should be placed on symptom interaction between the physician and patient, to ensure more optimal interpretation of the symptoms, as well as their severity.

The severity of signs and symptoms has often been misrepresented in most diagnostic models. Some past research has indicated that the symptoms and signs are dichotomous [26]. That is, the signs and symptoms either exist or do not exist, with no degree of severity distinguishable. This aspect definitely affects the preciseness and accuracy of these models, since physicians admit that the degree of severity of attributes influences their decisions. Other techniques used to represent the severity of signs and symptoms are somewhat artificially imposed [23]. Included in these techniques are assumptions that the severity of a symptom is normally distributed about a mean or fits some other continuous distribution. Others impose rigidly bound severity categories into which the symptom or sign must fall. But symptoms are felt and described by patients with varying levels of sensitivity and expression. Thus, their descriptions of these terms are too fuzzy and imprecise to be made to fit any distribution or limited to one of two or three mathematically imposed categories whose bounds can, at best, be vague.

Mathematical techniques used to model the preliminary diagnostic decision and the initial hypothesis are discussed in the

following section when final diagnostic decision models are enumerated and evaluated.

Diagnostic Reasoning and the Diagnoses

Once the history and signs have been collected and evaluated, explanations for these attributes should be developed by the physician via a diagnostic reasoning process. The initial preliminary diagnosis of the last subsystem, as well as revisions made in this subsystem, are used by the physician to guide steps taken in various phases of this reasoning process.

The process's initial phase usually involves further analysis of a specific aspect of the attribute, the location factor. The portion of the body that is the attribute's structural or functional source is defined to be its domain. This domain may encompass the location of more than one structural or functional attribute. Recognition of attributes with the same or related domains restricts investigation to a more limited physical area. The domain of each attribute can usually be classified into at least one of the following categories:

1. Organ
2. Region
3. Channel
4. System

An organ is a discrete structure, usually consisting of a cover, interior lining, and enclosing vessels and ducts. Examples include the skin, the brain, the stomach, the lung and the heart. A region is more generalized in location than an organ, and pertains to an anatomically defined part of the body, containing more than one organ.

This category comprises such areas as the abdomen, chest, head and pelvis. A channel consists of a group of structures or organs connected in a direct anatomic sequence for the functional purpose of transmitting a flow from one part of the body to another. The digestive tract, the nervous system, and the urinary tract are examples of these channels. The system is the last category and differs from the previous three types of domains in that it is defined according to its function rather than its location. For instance, the endocrine system regulates other systems and the cardiovascular-pulmonary system circulates oxygenated blood. If a more precise attribute location than the domain is known, then the location of the attribute can be classified as a focus within that domain. Naturally, attributes such as fatigue or weakness, which exhibit no specific location, cannot be placed into the above categories.

The next phase of the reasoning process involves determining the precise effect the causes of the attributes have on their domain. These causes may be revealed to a certain degree in the preliminary diagnosis of the physician. If there exists a gross abnormality of the domain, then a disorder exists. The degree of this disorder is determined from the signs, symptoms, severity of signs and symptoms, sign and symptom frequency, as well as clinical tests performed for this purpose. A disorder of the domain's structure is called a lesion, and, of function, a dysfunction. Factors influencing the importance of a lesion include size, composition, and location. Factors influencing dysfunctions are quantity of function, operation, and direction. It is unlikely that these factors can be observed during physical examination of the patient, unless the dysfunction or lesion occurs near the surface of an observable

domain. If information concerning the precise domain of the attributes, the degree of disorder, or other factors of the disorder are not known from the physician's observations, selected clinical or diagnostic tests may be performed to provide the needed information. Information obtained from these tests should be added to the attribute information and a revised preliminary diagnosis made. Aspects of this testing will be discussed further in the next subsystem.

Once the disorders and levels of disorders have been established, direct effort should be made by the physician to reach an effective final diagnosis. The last preliminary diagnosis may be of such a nature that no additional information is needed to decide upon the final diagnosis. However, if the preliminary diagnosis does not designate a specific cause of the attributes with an acceptable level of certainty, then other additional information is needed. At this point, the physician makes a differential diagnosis. In doing so, he selects a group of possible diseases or causes of the attributes from his final preliminary diagnosis. Through diagnostic tests or personal evaluation, the possible diseases are eliminated until the most likely disease candidate or candidates are known. Further tests, if available, can be performed for re-enforcement of the correct diagnosis.

The questionnaire revealed other eseential information about this subsystem and the state of the diagnosis. The following percentages indicated how often a sick patient's final diagnosis fell into the given categories:

Single Disease	62.8 percent
Multiple Disease	20.0 percent
A Few Possible Diseases	5.7 percent
No Precise Disease	5.7 percent
Only Known That Certain Diseases Do Not Exist	5.8 percent

Almost two-thirds of the diagnoses reached by physicians answering this survey limited the possible diagnostic candidates to a single disease. One-fifth of the diagnoses designated more than one disease present in the host. The remaining categories together accounted for a little more than 15 percent of the diagnoses. For one of these categories, the physician was not able to reach a diagnosis of a specific disease entity, so treatment was started aimed at a few possible diseases. The last two categories include diagnoses that are less precise than the first categories. In one of these, the attributes do not correspond to any specific disease entity, thus the disease is identified only by these attributes. For the other category, the diseases of concern have been eliminated from the possible candidates, leaving those lacking in enough severity to demand additional attention.

The questionnaire results also revealed how often physicians seek assistance from other physicians before reaching a diagnosis. Physicians consulted one another in approximately the same percentage of cases, regardless of the patient's location at the office or the hospital. No physicians consulted with other physicians for more than 30 percent of their cases and the majority consulted for less than ten percent. Thus, most of the diagnostic decisions in this subsystem are made by

the single physician attending the patient.

Certain aspects of this diagnostic subsystem have been modelled extensively, while others have received relatively little attention. Models involving the actual diagnosis decisions (preliminary and final) are abundant in decision theory and information science research. These models involve pattern classification techniques oriented towards computerized diagnosis and often requiring huge amounts of data not readily available. They are also limited to diagnosing single disease entities and cannot cope with multiple diseases in the same patient. Furthermore, the output of these models rarely determines the diseases' stage of development and assumes attribute consistency throughout the history of the disease in the patient.

The two techniques used in the vast majority of these models center on cluster analysis or Bayes' optimality criterion (formula). Application of Bayes' formula [31,39,48] requires a subjectively determined a priori probability of each disease, and conditional probabilities for specific symptoms and signs, given a certain disease is present. The objective is to find the disease with the maximum probability, given that a specific set of attributes exist. If this initial probability is not high enough, some sequential models allow additional information to be added until the probability reaches a certain acceptable level.

More complex forms of Bayes' equation have been applied when symptom and sign severity are assumed to be normally distributed. A multivariate normal distribution is then used in calculating the needed conditional probabilities. Additional complexities arise when

the appearance of symptoms and signs are dependent upon one another. The error of many Bayesian models is to assume statistical independence for the sake of computational simplicity [23]. But the interrelation of attribute appearance, location and severity is pertinent to an effective diagnosis. To alleviate this problem, complex convariance structures, requiring vast data bases or precise subjective probability guesses, have been suggested. But often attribute dependency is non-homogeneous throughout the severity of the attribute, making this approach incorrect. Another technique, involving the clustering of dependent attributes, has been developed to solve this problem [43]. In this method, signs and symptoms within the cluster are dependent upon one another, but the clusters themselves are independent. Probabilities are then based on the independent clusters rather than on individual attributes. Other techniques have been developed to handle these dependent attributes, including adjustments on solutions where dependency was assumed [42].

Perhaps the greatest drawback to Bayesian models representing the diagnostic decisions concerns the definitions of the probabilities themselves. In these models, the significance of each attribute to the disease is determined solely by the percentage of time it is present when the disease occurs. But in the actual decisions, this is not the only factor influencing the importance of an attribute to a disease. A symptom, such as fatigue or weakness, may occur with relatively high probability, but its importance to the diagnosis may be small. On the other hand, a symptom such as anginal pain or syncope may occur with relatively low probability, but its occurrence may be a cardinal

or key attribute of the disease. Thus a more fuzzy concept, such as the pertinence of the attribute to the disease, may be needed to represent attribute significance in the diagnosis decision.

Many of the difficulties associated with Bayesian models are also encountered in the use of cluster analysis techniques [23,26,34]. Cluster analysis techniques involve classifying a patient to the proper disease by determining differences between the patient's attribute vector and attribute vectors for each of the possible diseases. These models include classification according to the mean squared error criterion, closest neighbor criterion, and least Euclidean distance criterion. Most models of this type are oversimplified with respect to symptom and sign representation and do not follow the sequential nature of the process when modelling preliminary or final diagnoses. The significance of the attributes to each of the diseases is often omitted and improperly assumed to be constant for all attributes. Although past clustering techniques lack development and specificity, the possibility for more realistic models lies with the adaptability of these methods.

Unfortunately, few references include any modelling of the diagnostic reasoning process described earlier. The models just discussed make no attempt to explain the occurrence of attributes, but only classify them to reach a diagnosis. As stated by Leroy, "Traditional principles of science require that natural phenomenon be explained, not merely labelled. Scientific connection for relations between symptoms and diagnoses is performed by sequential explanations." [37]. Thus, additional modelling efforts need to be initiated

in the diagnostic reasoning area to alleviate this deficiency.

Tests for Additional Information

The purpose of the previous subsystems was to gain information, evaluate this information, and reach the proper diagnoses. At many points in this sequential process, improper descriptions, observations, and interpretations, as well as omission of information could result in an improper diagnosis. Thus, clinical and diagnostic tests are often performed throughout the diagnostic subsystem to gain the information needed to guarantee a proper diagnosis. These tests involve investigation of body fluids and tissue, as well as the use of electronically oriented machines that produce such evidence as X-rays or E.C.G. A variety of specific interrelated reasons exist for the use of diagnostic and clinical tests. Included in these are the following:

1. Development of information for determining the presence or degree of a disorder.
2. Qualification and quantification of the pertinent aspects of lesions and dysfunctions.
3. Elimination of a disease candidate from consideration during the differential diagnosis phase.
4. Validation of a disease or diseases as the proper final diagnosis.

Rather than perform a test haphazardly, the physician must determine if the need for the test outweighs its contraindicants. To determine this need, the physician must ask himself: "How important can the results of this test be in determining or changing the final diagnosis? Will the results of this test provide a better

understanding of the disease process in this patient? Will the patient benefit from what I am looking for if I find it?" Specific answers to these questions reveal the importance or significance of the test result to the diagnostic process being carried out.

Certain factors must be weighed against the test's need to determine if the performance of the test is justified. Unfortunately, these factors are somewhat fuzzy and difficult to quantify. The questionnaire revealed the following as determining factors:

1. The usefulness or the physician's past success with the test.
2. Danger of the test.
3. Cost of the test.
4. Convenience of the test to the patient and physician.
5. The test's discomfort or pain.

The first three factors are the most important to physicians selecting a test. Certainly the danger factor, as well as the cost factor, could overshadow the real need for the test. Another factor mentioned by the physicians included their familiarity and understanding of the test. This is somewhat incorporated into the physician's past success with the test, since he must be familiar with the test to determine its success rate.

Therefore, before a test is performed, it should be evaluated with respect to its need versus its usefulness (success), cost, convenience, danger, discomfort, and familiarity. If the need does not warrant the cost and other negative factors, then the test should not be performed. If this evaluation results in more than one

favorable test, then the best test should be selected from the beneficial candidates.

The following medical criteria are often used by the physician to determine the usefulness or success of these tests:

1. Sensitivity
2. Specificity
3. Predictive Value

The term sensitivity corresponds to the percentage of positive test results in patients actually possessing the disorder or disease the test was aimed at. This sensitivity may vary with the development of diseases or disorders, since the early stages are frequently undetectable. Specificity is derived from the percentage of negative results among people who do not have the tested disease or disorder. For a test to be useful, both the specificity and the sensitivity should be relatively high. The predictive value of a positive test result defines the percentage of positive results that are true positives. This value varies with the sensitivity, specificity, and incidence of the test. If the test is not performed often, then the values of specificity and sensitivity may possess a great degree of error. Thus, before considering a test of this type, the physician should determine whether the test sensitivity, specificity, and predictive values are adequate to provide clinically useful information.

Once the best possible test has been selected, the physician, laboratory assistant, nurse, or other person performs the test after possible preparations. The results of these clinical tests can be evaluated with respect to their "degree of abnormality." Test results

of this type include blood pressure, leukocyte level, and serum cholesterol level. By increasing above or decreasing below the "normal" range, the degree of abnormality of the test result increases. On the other hand, many tests, primarily those aimed at the final diagnosis, are concerned more with dichotomous results. Tests of this type either prove or rule out specific patient abnormalities.

Mathematical models for determining the best possible test at a specific time in the diagnostic process have been developed in past research. The two methods found while studying for this research include the use of decision trees, as well as a single test selection function. The decision tree approach [32] is usually incorporated in a model of the whole decision process, while the test selection function represents only this subsystem. Models of the decision tree type do not seem sensitive to all the factors which influence test selection. Also, sequential testing cannot be easily handled by this decision tree approach without re-evaluation of the tree at each step. The test selection itself is often based solely on a priori probabilities, since the overall analysis usually starts with test selection, followed by diagnosis and treatment selection. Other difficulties, normally encountered when using decision trees, are studied later in treatment modelling analysis.

Unlike the decision tree approach, test selection by a single function seems more appropriate and representative of the process's important factors [41]. Sensitivity, specificity, predictive value, and other variables can be incorporated into this test selection

function. Sequential testing can also be evaluated with less difficulty, since components of each test function usually remain specific and constant for a given test. Consequently, future research needs to be performed in this area to better illustrate the assets of this method.

The General Treatment Process

The treatment process begins when the final diagnosis is reached in the diagnostic subsystem. Once the physician has explained the incidence of signs, symptoms, and other attributes and makes this diagnosis, he begins to consider a treatment or therapy for the sickness. His treatment decisions normally fall into two classifications - therapeutic and environmental. Therapeutic decisions involve the selection of the treatment, while environmental decisions are aimed at the management of the host during treatment administration. The treatment process may last only a short period of time or the rest of the patient's life. The length of the process depends upon the effectiveness of the treatment and the nature of the disease or disorder present.

Selection of the Treatment: Therapeutic Decisions

The category of the final diagnosis directly influences the treatment strategy as well as the treatment selection itself. If the physician has determined conclusively that a particular single disease is present, the best possible treatment for that disease is selected. If the physician has designated multiple diseases in the diagnosis, then treatment should be aimed at each disease present. Unfortunately, one disease' treatment may interfere with the treatment of another, creating complications. In this case, the most serious or dangerous

disease should be selected and treated initially. When the diagnostic process names no specific disease with certainty, then one or more diseases of a finite set of possibilities may exist. The usual strategy would be to select the treatment for the most likely disease. But this treatment could be dangerous to one of the other contingent diseases if the most likely candidate does not subsist in the patient. Thus treatment consequences must be evaluated in the selection of the proper remedy when the diagnosis falls into this category. When no precise disease can be designated to explain the attributes, then treatment must be based only on these attributes. For sicknesses where the physician knows what diseases do not exist, the serious diseases have been eliminated from the list of possible candidates. In these instances, no treatment or only remedial treatment may be necessary. This concludes the strategies for the given diagnosis categories.

The physician must use this strategy to select a treatment from one of the following therapeutic categories:

1. Diet
2. Drugs
3. Surgery
4. Radiotherapy
5. Physiotherapy or Physical Therapy
6. Psychotherapy
7. Combination of Categories
8. No Specific Treatment

The idea behind the treatment process is to repeat, reproduce, and surpass past success in remedying similar diagnostic situations.

Only rarely is the treatment process creative in nature, since certain moral implications as well as legal problems tend to discourage this. To select the proper treatment, the physician can study the past etiology and pathogenesis of the diseases specified in the diagnosis to determine the best medical treatment. Physicians occasionally lack information about the patient or the treatment's effect on the patient. Thus, clinical tests must be run to determine the appropriateness of treatment for a specific patient.

If more than one treatment is deemed medically effective and appropriate, then other factors may influence the selection of the best treatment. These factors, somewhat fuzzy in nature, are listed below along with their average occurrence in the treatment decision.

Factor	Occurrence
1. Danger of the treatment	Often
2. Extreme cost of the treatment	Often/Occasionally
3. Desirability of treatment to patient	Occasionally
4. Pain of the treatment	Occasionally
5. Physician's success with the treatment	Occasionally
6. Time consumption of the treatment	Occasionally/Seldom
7. Unfamiliarity of treatment	Occasionally/Seldom

These influences, as well as the predicted medical effectiveness of the treatment, should be weighed against the patient's need for the treatment. Criteria for determining this treatment effectiveness and need are so similar to those discussed in the test selection subsystem that they will be omitted from this phase of the research. Once these

factors are evaluated with respect to one another, the proper treatment is selected via the treatment strategy.

Models for analyzing the selection of treatment have been developed in detail. These models use, almost exclusively, a decision tree approach to the problem [20,32,44]. They are invariably applied to problems where some uncertainty exists as to what disease is present. This would correspond to the category of "a few possible diseases" in the diagnostic subsystem. When using the decision tree approach, the selection of treatment depends upon the following information:

1. A priori probabilities of the diseases being present.
2. Conditional probabilities of treatment results.
3. Conditional consequences or utilities given the treatment and diseases.

While the structure of decision trees may assist the physician in analyzing his alternatives, the quantitative measures cannot hope to be very exact. By the time all the conditional probabilities for a very large tree have been determined, variations in quantification could give an incorrect treatment selection. The consequences or utilities also may vary between physicians since personal ethics and experience are subjectively determined by each physician's own personal criteria. Any pruning of large decision trees, as suggested in past research, may be dangerous since alternatives with very low probabilities could have drastic consequences, demanding additional consideration. Therefore, the decision tree is a good technique for enumerating treatment alternatives and consequences, but large data bases may be needed before quantitative consistency is attained.

Another model has been developed to handle a very specialized case of the treatment selection process. Developed by Knapp and Oeltjen [35], this method uses a benefit-risk ratio to select drugs which have potential side affects as well as potential risks. Risk interpretation by different physicians is questioned and believed to vary somewhat, as was the case in evaluating the utilities used for the decision trees. To validate this hypothesis, diseases were graded with respect to seriousness. Then physicians weighted each drug, using a risk versus need or benefit criterion. The experimental result verified that interpretation of risks and benefits varied between physicians.

Treatment of the Patient-Environmental Decisions

Once the treatment or therapy has been selected, the physician explains, prescribes or performs the treatment of choice. The physician interacts with the patient, often altering aspects of the treatment to meet the patient's needs and desires. In many instances, the physician must explain the treatment, the appropriateness of the treatment, as well as fuzzy and nonfuzzy instructions for administration of the treatment. Decisions made during this process are considered environmental decisions since they involve the management and manipulation of the patient. If the treatment involves diet, then the patient is instructed in the low fat, low sodium, or other specialized diet. When drugs are prescribed, the physician explains their use as well as their possible hazards or risk factors. The patient should then follow these instructions and observe possible adverse or negative reactions to the drugs. If the treatment involves surgery, then the

patient should be aware of the surgical process and have knowledge of its limitations, hazards, and consequences. During radiotherapy, the patient has to be advised of the side affects of the treatment as well as the degree of possible success. Psychotherapy is a totally interactive, questioning process between physician and patient, often involving repeated visits, since therapy is performed via verbal communication. For some treatments, success or failure may be fairly easy to determine. But with psychotherapy, the results may be difficult to classify into such discrete categories. In this age of abundant medications, explanation, and justification for offering no treatment may be difficult, especially if the patient has determined he needs some sort of relief. Thus, each of these various types of treatment require uniquely different environmental decisions of the physician.

If time, test results, or continued symptoms and signs reveal that the treatment offered is not successful, then additional actions and decisions must be made. If the treatment selected was actually the best therapeutic treatment, then several reasons could account for the unsuccessful result. First of all, the instructions given the patient concerning self administration of drugs, physical therapy, or diet could have been imprecise, incomplete, or too complex for the patient's comprehension. The patient may have been unattentive, uncooperative, or apathetic towards listening or abiding by the treatment instructions. Also, an effective treatment for one group of patients may be ineffective for another group, since certain aspects of the patients' past history and body chemistry vary. An error in

the actual treatment process, not noticed during treatment application, may result in less than optimal results. For instance, psychotherapy or specific surgery may be the treatment of choice, but if it is administered incorrectly, adverse consequences could occur.

A treatment could also be ineffective if the physician's treatment selection was not actually the best possible choice. This could be due to lack of critical information or just poor judgment. If the treatment selection is improper, then either the selection was wrong or the diagnosis, on which it was based, was incorrect. Information gained while administering the erroneous treatment should give insight into this problem. If only the treatment selection was inappropriate, then a new treatment should be selected by returning to the previous treatment subsystem. If the diagnosis was wrong, then a new final diagnosis must be reached via the diagnostic reasoning process, which should incorporate any additional information gained. In a general systems theory sense, the diagnostic and treatment process could thus be cyclic in nature [45], if ineffective treatments and diagnoses continued to be selected.

Constraints by the Environment

Constraints by the environment encompass a variety of factors which evolve outside the physician-patient relationship, but influence it directly. Many of these factors act as limitations or guidelines for physician oriented decisions in the relationship. Such constraints are derived from the physicians' professional establishment, legal and governmental sources, private business, and community resources.

Somewhat ill-defined restrictions, involving professional ethics, are administered primarily by legal and professional establishments. These legal and professional standards enumerate and assess specific instances of improper practice and poor decision making. Professional criticism may include informal conversation or written comments and guidance by fellow physicians through prominent publications. More specific are the legal constraints which are imposed by such agencies as the Food and Drug Administration. They tend to deal mainly with prohibiting the performance of certain tests and treatments. Certain techniques and drugs, deemed unsatisfactory due to possible dangerous consequences to patients, have been banned from therapeutic consideration in this country. Other related factors include varying degrees of control on the fiscal policies of physicians. This involves private insurance corporations and government programs such as Medicare and Medicaid, which place bounds or restraints on the patient-physician fiscal relationship. Due to the influx of medical malpractice suits and increased costs, insurance companies also influence the medically oriented decisions made by physicians. Such influence encourages incidences of overtesting in reaching a diagnosis and more conservative treatment selection.

A community's attitudes and medical resources also affect many of the actions and decisions made by local physicians. Abstracted from its population, the community as a whole imposes ethical, practical, and medical expectations on its physicians. This is demonstrated by the lack of patients and financial support from a community that desires alternate medical care. Another community constraint

involves medical facilities which are substandard in quality and quantity. As mentioned in a previous subsystem, inconveniences of this type curb the number of possible tests under consideration and eliminate possible treatments from further evaluation.

CHAPTER III

THE DIAGNOSTIC DECISION PROCESS: A GENERAL PROBLEM FORMULATION

Since the decision processes analyzed and discussed in the previous chapter are complex and abound, emphasis on mathematical models incorporating fuzzy set theory has been restricted to specific diagnostic decisions. These decisions involve the use of diagnostic information and the reaching of effective medical hypotheses, preliminary, and final diagnoses. The primary reason for the selection of this decision process centers on the fact that other diagnostic decision processes, such as the test selection process and the diagnostic reasoning process, hinge on decisions made in the preliminary diagnoses and are thus somewhat secondary. An effective general model for reaching the various diagnoses must be developed first, so that these secondary decision processes might then be integrated into a comprehensive overall model of the general diagnostic decision process. Furthermore, development of models for treatment decisions would be inappropriate at this early modelling stage, since these decisions would depend on medical and mathematical information evaluated during the diagnostic process. As a result, the mathematical modelling illustrated in this chapter pertains to the gathering and evaluating of patient information and the incorporation of this information into a general diagnosis making model.

Consequently, the prominent model of this research deals with the decision upon a medical hypothesis, preliminary diagnosis, or final diagnosis. Fortunately, the nature and structure of these diagnostic decisions are similar, only the information needed is different. There are numerous classical techniques used to model these decisions, many of which were mentioned and evaluated in the previous chapter. The most formidable classification techniques applicable to this modelling include Bayesian statistics and cluster analysis. The shortcomings of Bayesian techniques, the most prominent method of previous models, were detailed in Chapter II. Thus the techniques of cluster analysis have been deemed the most satisfactory for this modelling phase. The specifics of these techniques will be delayed until mathematical representations of patient information, needed for these decisions, are presented.

Patient Information: Mathematical Representations

The modelling of the physician's evaluation of patient information is given initially, since its mathematical structure is pertinent to the formulation of the general diagnosis decision. The information under consideration, discussed in Chapter II, can be classified into one of the following categories:

1. Past patient history.
2. Present symptoms.
3. Signs observed upon physical examination.
4. Results of clinical and diagnostic tests.

The various diagnosis decisions are usually based on more than one of

these categories. A medical hypothesis requires information from the first two categories, an initial preliminary diagnosis requires information from the first three categories, and additional preliminary diagnoses and the final diagnosis usually require information from all these categories. The evaluation of information from each of the preceding categories has been mathematically represented in the subsequent sections.

Past Patient History

As mentioned previously, certain specific aspects of the patient's past history often aid in suggesting possible candidates for preliminary diagnoses. Such nonmedical aspects may include an occupation in a dangerous chemical environment, the age group with the greatest disease susceptibility, or membership in a high incidence group of the population. Medical aspects include the patient's past diagnosed diseases and disorders, as well as symptoms that were not diagnosed or observed by a physician at the time of their occurrence.

Conventional modelling techniques represent this past history in a variety of ways. In Bayesian methods, the physician assigns a subjective a priori probability, $p(i)$, for each considered disease i, which supposedly reflects this past history. Other methods may omit this information entirely. This research attempts to represent past history with more clarity and detail, in an effort to remedy the incomplete use of information in the above techniques.

Let N be the set of diseases under consideration for the diagnoses. For each disease $i \in N$, there exists a finite set, Ω_i, of prominent history aspects, similar to those mentioned above. This set

includes nonmedical aspects as well as medical aspects which are physician designated diseases or disorders. Let

$$\Omega = \bigcup_{i=1}^{n} \Omega_i \tag{1}$$

where

n = number of elements in N.

This set Ω has as its elements all aspects of patient history relevant to the diagnoses. Each of these aspects or elements of Ω is either present or absent from a particular patient's history and is assumed to be binary. To represent this, a 1 x m matrix H is created for each patient, such that each element of this matrix relates the presence or absence of the corresponding element of the set Ω. Thus

$$H = [h(1) \quad h(2) \ \ldots \ h(m)] \tag{2}$$

where

$h(i) \in I$

$I = \{0,1\}$

m = number of elements in Ω.

Complexities of this modelling arise when physicians encounter incomplete designations of medical aspects in the patient's history. For instance, rheumatic fever or syphilis may be important to the diagnoses of valvular heart diseases, but frequently the patient may have had one of the diseases and not had it properly diagnosed and

designated. Thus the physician must often determine the presence or absence of rheumatic fever or syphilis based on past, undiagnosed symptoms. To mathematically convey this, let $\Omega_A \subset \Omega$ such that the p elements of Ω_A correspond to the p diseases or disorders which might be missed or go undiagnosed, where $p \leq m$. Consider the p elements of Ω_A and their corresponding elements in the matrix H. Since both the set and matrix are finite, the matrix H can be ordered or restructured so that its first p entries correspond to the elements of Ω_A. Thus

$$H = [h(1) \quad h(2) \ \ldots \ h(p-1) \quad h(p) \quad h(p+1) \ \ldots \ h(m)] \ . \quad (3)$$

From this matrix H, submatrices H_A and H_B can be formed such that $H = [H_A, H_B]$, where

$$H_A = [h(1) \quad h(2) \ \ldots \ h(p)] \ , \quad (4)$$

$$H_B = [h(p+1) \quad h(p+2) \ \ldots \ h(m)] \ .$$

The presence or absence of each history aspect in Ω_A, represented in the binary matrix H_A, might then be determined from past, undiagnosed symptoms and sicknesses. Since not all these past symptoms will be recalled by the patient, the physician only considers the "prominent" symptoms of past sicknesses. This concept of prominence is somewhat fuzzy in nature and can be quantified via fuzzy set theory. Let β_j be the set of possible symptoms for past undiagnosed disease j, where $1 \leq j \leq p$. The fuzzy set θ_j, containing the "prominent" symptoms of past disease j is defined as

$$\theta_j = \{(x_{ij}, u_{\theta_j}(x_{ij})) \mid x_{ij} \in \beta_j, u_{\theta_j}(x_{ij}) \geq \alpha\} \tag{5}$$

where

$$u_{\theta_j}(x_{ij}) \in [0,1]$$

$$0 \leq \alpha \leq 1$$

Whenever the physician designates the membership function $u_{\theta_j}(x_{ij})$ of symptom x_{ij} for disease j over a specified level α, the symptom of β_j becomes a "prominent" symptom. For each fuzzy set θ_j, $1 \leq j \leq p$, construct a matrix V(j), where the elements of V(j) represent the presence or absence of x_{ij}, $\forall\ x_{ij} \in \theta_j$.

$$V(j) = [v(1,j) \quad v(2,j) \quad \ldots \quad v(k_j,j)] \tag{6}$$

where

$$v(i,j) \in I$$

$$I = \{0,1\}$$

$$i = 1,2,\ldots,k_j$$

$$1 \leq j \leq p$$

k_j = number of elements in θ_j

The representation, v(i,j), for symptom i of disease j, is binary and thus assumed to lack severity levels since the specifics needed for severity determination are forgotten or altered with the passing of time.

If disease j is not designated in the patient's history, but symptoms of disease j have existed in the patient's past, then the matrix V(j) must be determined. This matrix V(j) is then used to complete the patient history matrix H. For a specific patient and undiagnosed disease or disorder j,

$$V(j) \xrightarrow{f(V(j))} h(j) \qquad (7)$$

such that

$$f(V(j)) \in I$$

$$I = \{0,1\}$$

$$1 \leq j \leq p$$

Therefore, a vector of symptoms V(j) is mapped (nonfuzzy) into the jth element of H_A, such that h(j) equals 0 or 1.

The function f(V(j)) may be very simplistic in nature, or may be similar to the function presented as the general diagnoses decision model. The number of elements for the sets Ω_A and θ_j, $1 \leq j \leq p$, given as examples later in this research, are relatively few. Thus, the simplistic approach to this function, taken for this research, is presented in the illustrative example of Chapter IV. Using this mapping, the past history matrix H_A can be quantified for all possible elements h(j), $1 \leq j \leq p$. The completed matrix H is then used later when the various diagnoses decisions are modelled.

Present Symptoms

The interpretation of patient symptoms, as well as the evaluation

of past modelling representations, was discussed at length in Chapter II. Each of the past models studied begins the decision process with the symptoms already designated or severity levels determined. As was mentioned previously, these severity levels are often incorrectly assumed to always be binary. Since subsequent modelling of this research attempts to determine the presence of a disease as well as its stage of development, the preciseness of these symptom severity levels becomes increasingly significant. Medical designation of these levels should not be given outright, as in previous models, but should come directly from information related by the patient concerning his present medical problems. By using the latter approach, the severity levels of symptoms can be more precisely defined and designated.

To begin this phase of modelling, let Π_i equal the set of possible problems that might be observed or experienced by a patient with disease i, where $i \in N$.

Define

$$\Pi = \bigcup_{i=1}^{n} \Pi_i \tag{8}$$

where

$$n = \text{number of elements in } N.$$

This set Π comprises the patient related problems that are encountered for all diseases under consideration. The elements of this set are not medically designated symptoms, but are patient descriptions such as dizziness, chest pain, and inability to breathe properly. Associated

with each of these problems q, $q \in \Pi$, is a set of factors Δq, which are important to the medical designation and severity determination of problem q. Elements of Δq make up what was called in Chapter II the undesignated symptom profile or problem profile. Each problem profile set Δq may have the following subsets of discrete information:

$$\begin{aligned}
\Delta q1 &= \{\text{location of problem q}\} \\
\Delta q2 &= \{\text{longevity of problem q}\} \\
\Delta q3 &= \{\text{continuity of problem q}\} \\
\Delta q4 &= \{\text{defining aspects of intermittent problem q}\} \\
\Delta q5 &= \{\text{specifics for severity determination of problem q - fuzzy descriptions}\}
\end{aligned} \tag{9}$$

where

$$\Delta q = \Delta q1 \cup \Delta q2 \cup \Delta q3 \cup \Delta q4 \cup \Delta q5$$

Let

$$\Delta = \bigcup_{q=1}^{r} \Delta q \tag{10}$$

where

$$r = \text{number of problems in set } \Pi$$

If β is the collective set of medically designated symptoms for the diseases under consideration, then

$$\Pi \xrightarrow{f(\Pi,\Delta)} \beta \tag{11}$$

When determining the presence or absence of a medically designated symptom of β, this mapping is usually nonfuzzy in nature. But when the severity level of an element of β must be determined, then fuzzy sets are often involved in the mapping.

At this point in the modelling, matrix representations for the sets under discussion are introduced. Information obtained from the patient with regard to his problems, Π, and problem profiles Δq are incorporated into the matrix B, defined as

$$B = \begin{bmatrix} b(1,1) & b(1,2) & \dots & b(1,r) \\ b(2,1) & b(2,2) & \dots & b(2,r) \\ \vdots & \vdots & & \vdots \\ b(s,1) & b(s,2) & \dots & b(s,r) \end{bmatrix} \tag{12}$$

An entry of this matrix is denoted by

$$b(i,j) \tag{13}$$

where

$i \in I$

$j = 1,2,\dots,r$

I = the index for characteristics of the problem profile

$r = \max I$

For example, the ith row of this matrix might correspond to the location of problems j, $j = 1,2,\dots,r$. The jth column of this matrix corresponds

to the profile factors in the set Δq where $j = q$.

This matrix B should contain all the pertinent information obtained from the verbal physician-patient interaction during the initial interview. Once the information for B has been gained, the physician must medically designate the symptoms as well as the severity levels. From the set β of all possible medically designated symptoms under consideration, construct a matrix A such that

$$A = [a(1)\ a(2)\ \ldots\ a(t)] \tag{14}$$

where

$$a(i) \in [0,1]$$

t = the number of elements in β.

The variable $a(i)$ of A represents the severity of symptom i, with $i \in \beta$. If a symptom i is assumed to be dichotomous for the diseases under consideration, then

$$a(i) \in \{0,1\} \tag{15}$$

with

$$a(i) = \begin{cases} 0, & \text{if symptom i is not related by the patient} \\ 1, & \text{otherwise} \end{cases}$$

Otherwise, the severity of the symptom is considered to be pertinent information and can be represented as $a(i) \in [0,1]$. As $a(i)$ approaches 1, the severity of symptom i increases, while the severity of symptom i decreases as $a(i)$ approaches 0.

The previous mapping (11) of patient-related problem to medically designated symptoms can now be more precisely written in the matrix form as

$$B \xrightarrow{\delta(B)} A \tag{16}$$

In this matrix form one problem j may be mapped into one or more entries of A. Let $\bar{b}(j)$ denote a column vector of matrix B. As stated previously, this vector represents the problem j and the problem profile set Δj. The above mapping (15) can now be more specifically written as

$$\bar{b}(j) \xrightarrow{\delta(\bar{b}(j))} a(i) \tag{17}$$

where

$$j = 1,2,\ldots,r$$

$$i = 1,2,3,\ldots,t$$

When a(i) is binary, then $\delta(\bar{b}(j)) \in \{0,1\}$. For this case, the function δ is very simplistic and usually incorporates the presence or absence of a few factors, primarily those of location, in the problem profile.

If a(i) is not assumed to be binary, then more complex mappings involving fuzzy set theory exist. In this case, a component of $\bar{b}(j)$, say b(i,j), often represents the membership function for a fuzzy description of the severity of problem j. In this case

$$u_A(c) = b(i,j) \tag{18}$$

where

c = relative measure or description of the symptom severity.

A = fuzzy set evaluating severity descriptions.

The membership function may reflect the painfulness or blueness of symptoms such as headache and cyanosis. Fuzzy sets such as these are pertinent to determination of symptom severity levels. Thus for this problem case, the function of the mapping in equation (17) becomes

$$\delta(\bar{b}(j)) = \delta[b(1,j)\ \ b(2,j)\ \ldots\ u(c)\ \ldots\ b(k,j)]^T . \qquad (19)$$

This mapping can thus involve functions of fuzzy and nonfuzzy sets. Many times, only the membership function of a fuzzy set, $u_A(c)$, is needed to determine the severity of a symptom. In the example in the next chapter, the function frequently is

$$\delta(\bar{b}(j)) = u_A(c) \qquad (20)$$

or

$$\delta(\bar{b}(j)) = g(u_A(c))$$

where g is a simple function of fuzzy sets and other problem factors. Thus, fuzzy set theory is very useful in determining the severity levels of non-binary symptoms. Detailed examples will be given in Chapter IV to clarify this theory and to illustrate the specific fuzzy sets, their membership functions, and their corresponding functions $\delta(\bar{b}(j))$.

Signs Observed Upon Physical Examination

Signs are observed, evaluated, and designated by the physician upon physical examination of the patient. Upon careful inspection, the physician determines which signs are present or absent, as well as the signs' levels of severity. The modelling difficulties arise in determining the severity of the sign in such a way as to correspond

to other modelling in this research. Past or previous modelling representations of signs are the same as for symptoms. Thus, most of these models assume signs are dichotomous, while relatively few reflect degrees of severity. This section of research attempts correct this deficiency by more precisely modelling varying levels of sign severity.

For the diseases under consideration, let

$$\Phi_j = \{\text{signs of disease } j\}$$

where

$$j \in N$$

and

$$\Phi = \bigcup_{j=1}^{n} \Phi_j$$

Consider now the use of matrix notation to represent this information. Suppose

$$S = [s(1)\ s(2)\ \ldots\ s(f)] \tag{22}$$

where

$$f = \text{number of elements in } \Phi$$

$$s(j) \in [0,1] \qquad j = 1,2,\ldots,f$$

Thus, $s(j)$ reflects the severity of sign j in the patient. If the sign corresponding to $s(j)$ is assumed to be binary then $s(j) \in \{0,1\}$.

If severity levels of the sign are pertinent to the various diagnoses, then s(j) is nonbinary and $s(j) \in [0,1]$. For each nonbinary sign j, there exists a finite set of observable factors or observables, which assist in determining the level of severity for the specific sign. These observables make up what was termed to sign profile in Chapter II. Included in these observables are the locations of the signs and certain severity specifics. Rather than have a physician subjectively determine the level of severity of sign j, s(j), the use of a model incorporating these observables to attain this is proposed.

Let D denote a matrix of observables, such that

$$D = \begin{bmatrix} d(1,1) & d(1,2) & \cdots & d(1,f) \\ d(2,1) & d(2,2) & \cdots & d(2,f) \\ \vdots & & & \vdots \\ d(e,1) & d(e,2) & \cdots & d(e,f) \end{bmatrix} \tag{23}$$

with

e = number of elements in largest set of observables.

An entry in this matrix is denoted by

$$d(i,j) \tag{24}$$

where

$i \in I$, $j = 1,2,\ldots,f$, and I = index for each set of observables.

Each column of matrix D, say $\bar{d}(j)$, corresponds to the set of observables for sign j. To obtain the severity of sign j,

$$\bar{d}(j) \xrightarrow{\psi(\bar{d}(j))} s(j) \ . \qquad (25)$$

This mapping, as was the case in determining symptom severity, may or may not involve fuzzy sets. The function ψ is usually very similar in nature to the function δ used to map patient information into symptoms in equation (17). Nonfuzzy mappings exist when only location factors are pertinent to severity determination. If other factors exist, they often involve the use of fuzzy sets. For example, severity aspects of a systolic heart murmur may involve fuzzy sets with regard to the loudness and quality of the sound. Specific functional relationships of $\psi(\bar{d}(j))$ are omitted since the general structures are similar to those presented in equations (19) and (20) in the symptom modelling section.

Results of Clinical and Diagnostic Tests

The information gained from clinical and diagnostic tests is of extreme importance to the final diagnoses. This phase of research attempts to represent test results in the same framework as previous information, in an effort to incorporate them into the subsequent diagnosis model. Very few if any of previous diagnosis models have ever attempted to incorporate results from tests in their structure. The following model is presented in an effort to remedy this situation.

Consider the set, T_t, of possible test results of test t. This set T_t may be a set of discrete or continuous elements. Fuzzy set theory is introduced in an attempt to transform these test results

into a proper perspective and scale. For each test t, a fuzzy set Γ_t is created to represent the "abnormality" of the possible test results, T_t. The membership function, reflecting the degree of abnormality, must be determined for each set Γ_t, $t = 1,2,\ldots,k$, and k equals the number of tests performed on the patient. Let $r_t \in T_t$. The degree of abnormality of this test result is reflected by the membership function

$$u_{\Gamma_t}(r_t) \tag{26}$$

where

$$r_t \xrightarrow{u_{\Gamma_t}(r_t)} z(t)$$

and

$$z(t) \in [0,1]$$

For the k tests performed on a given patient, let

$$Z = [z(1)\ z(2)\ \ldots\ z(k)] \tag{27}$$

represent the test results evaluated via fuzzy set theory. Any test result can be mapped into the [0,1] continuum using this theory, so k is the number of tests performed.

For a certain group of tests, the results end in a single quantified measure, making it less difficult to determine the appropriate membership function. For example, the test performed to determine the serum cholesterol level of a patient results in a reading expressed in milligrams per 100 milliliters of serum. The degree of membership is thus a function of a single variable, where

$$u_{\Gamma_t}(r_t) = \begin{cases} 0 & , \text{ for } r_t < 260 \\ \frac{r_t}{340} - \frac{26}{34} & , \text{ for } 260 \leq r_t \leq 600 \\ 1 & , \text{ for } r_t > 600 \end{cases} \tag{28}$$

would be appropriate linear function for the fuzzy set of "abnormal" cholesterol levels. Similarly, many other membership functions for specific tests may be derived from single measurable results of the tests.

Other test results are not as simple, since more than one aspect is needed to determine their degrees of abnormality. The result of test t, r_t, may possess n defining aspects, so that

$$r_t = \{r_{t_1}, r_{t_2}, r_{t_3}, \ldots, r_{t_n}\} \tag{29}$$

where each r_{t_j}, $j = 1,2,\ldots,n$, is a defining aspect. Thus, the fuzzy mapping in equation (26) becomes

$$r_t \xrightarrow{u_{\Gamma_t}(r_{t_1}, r_{t_2}, \ldots, r_{t_n})} z(t) \quad , \tag{30}$$

where u_{r_t} is a membership function mapping $r_{t_1}, r_{t_2}, \ldots, r_{t_n}$ into [0,1]. Since these aspects, r_{t_j}, $j = 1,2,\ldots,n$, are very specific to the individual test t, generalities concerning the possible structure of $u_{\Gamma_t}(r_{t_1}, r_{t_2}, \ldots, r_{t_n})$ are difficult to make. Consequently, this will be investigated in a subsequent effort.

Summary of Mathematical Representations

The previous sections have given mathematical representations for each category of information needed in the various diagnoses decisions. A summary of the final representations, needed in the next modelling stage, is now presented.

The past patient history is represented by a binary 1 x m matrix H where

$$H = [h(1) \quad h(2) \ \ldots \ h(m)] \tag{31}$$

and m = total number of designated aspects of past history relevant to the diseases under consideration.

Furthermore

$$h(i) \in \{0,1\} \tag{32}$$

with

$$h(i) = \begin{cases} 1, \text{ if history aspect i, is present} \\ 0, \text{ otherwise} \end{cases}$$

The 1 x t matrix A represents all the medically designated symptoms for the possible diseases.

$$A = [a(1) \quad a(2) \ \ldots \ a(t)] \tag{33}$$

such that

$$a(i) \in [0,1]$$

t = total number of designated symptoms for all diseases.

For the signs observed by the physician, the 1 x f matrix S represents the signs under consideration.

$$S = [s(1) \quad s(2) \ \ldots \ s(f)] \tag{34}$$

where

f = total number of possible signs.

and

$$s(i) \in [0,1]$$

The last information category involves results of clinical and diagnostic tests. The final mathematical representation of these results is designated by the 1 x k matrix Z.

$$Z = [z(1) \quad z(2) \ \ldots \ z(k)] \tag{35}$$

where

k = number of tests performed on patient

and

$$z(i) \in [0,1]$$

For the various diagnoses under consideration, the following information matrices are needed:

Medical Hypothesis - {H,A}

Initial Preliminary Diagnosis - {H,A,S}

Other Preliminary Diagnosis - {H,A,S,Z}

and

Final Diagnosis - {H,A,S,Z}

Note that the matrix Z increases in size as additional tests are performed. The general diagnosis model incorporating these matrices of information in specific diagnosis decisions is developed in the subsequent section.

The Diagnosis Decision Model

General

The objective of the diagnosis decision model is to assign a patient, possessing information matrices H,A,S or Z to a stage of a specific disease or a group of diseases. These matrices represent the medically designated stage space of the patient. The earlier preliminary diagnoses tend to be aimed at selecting a group of possible diseases or eliminating diseases from consideration. The later diagnoses are aimed more at naming a single or a few diseases, existing at a specific stage of development. For any of these diagnoses, the information matrices of the patient must be compared to similar matrices for the stages of the possible diseases. The patient is grouped or clustered with the stage of disease which is the "closest" or "most similar". As was stated earlier, techniques of cluster analysis are used to mathematically analyze this patient - disease stage comparison.

The patient's history matrix, H, must be compared to the binary history matrix H(i) of each disease i. The history matrix for a given disease is constant, regardless of the disease's stage of development. Thus

$$H(i) = [h(i,1)\ \ h(i,2) \ \ldots \ h(i,m)] \tag{36}$$

where

$$h(i,k) \in \{0,1\}, \qquad i \in N$$

and

$$m = \text{number of history aspects.}$$

The element h(i,k) takes on a value of 1 if history aspect k has ever influenced the occurrence of disease i in the patient, and 0 if not.

The matrix A of present patient symptoms has to be compared to each symptom matrix of disease i at its development stages. In order to simplify the notation in this and the following sections, the finite number of possible diseases i at their finite number of development stages are numbered sequentially, j = 1,2,...,g, where g equals the total number of development stages for all diseases under consideration. The matrix A(j) thus contains the symptom severity specifics needed to reflect disease stage j. This matrix numerically designates the upper and lower bounds of the symptom's normal range of severity for a given stage. Thus

$$A(j) = \begin{bmatrix} alb(j,1) & alb(j,2) & \dots & alb(j,t) \\ aub(j,1) & aub(j,2) & \dots & aub(j,t) \end{bmatrix} \tag{37}$$

where

alb(j,k) = lower bound of kth symptom for disease stage j,

aub(j,k) = upper bound of kth symptom for disease stage j,

t = number of possible symptoms,

$alb(j,k) \in [0,1]$

$aub(j,k) \in [0,1]$

and

$$alb(j,k) \leq aub(j,k)$$

When symptom k is assumed dichotomous or binary, then

$$alb(j,k) = aub(j,k)$$
$$alb(j,k) \in \{0,1\} \tag{38}$$

If symptom k is not dichotomous, then an interval of severity may exist for the symptom at a disease stage j. The bounds of this interval are derived from the same specifics used to determine a patient's symptom severity, a(k).

A patient's information matrix, S, of observed signs, must be similarly compared to sign matrices S(j) for disease stage j. The severity levels of signs are similar to those of symptoms. Thus

$$S(j) = \begin{bmatrix} slb(j,1) & slb(j,2) \ldots slb(j,f) \\ sub(j,1) & sub(j,2) \ldots sub(j,f) \end{bmatrix} \tag{39}$$

where

$slb(j,k)$ = lower bound of kth sign for disease stage j,

$sub(j,k)$ = upper bound of kth sign for disease stage j,

f = number of possible signs,

$slb(j,k) \in [0,1]$

$sub(j,k) \in [0,1]$

$slb(j,k) \leq sub(j,k)$

As was the case with symptoms, a dichotomous sign k is assumed to have a constant level of severity, such that

$$slb(j,k) = sub(j,k)$$
$$slb(j,k) \in \{0,1\} \tag{40}$$

If the sign k is not dichotomous, then a interval of severity may exist for the sign at disease stage j. The bounds of this interval are derived from the same specifics used to determine a patient's sign severity, s(k).

The expandible matrix Z of patient test results has to be compared to the expected range of test results for a disease at a given stage. As was with symptoms and signs, a lower and upper bound for the evaluated test results must be known for each disease stage.

Let

$$Z(j) = \begin{bmatrix} zlb(j,1), zlb(j,2)\ldots zlb(j,k) \\ zub(j,1)\ zub(j,2)\ldots zub(j,k) \end{bmatrix} \tag{41}$$

where

$zlb(j,t)$ = expected lower bound of test result t for disease stage j,

$zub(j,t)$ = expected upper bound for test result t for disease stage j,

k = number of tests performed on the patient,

$zlb(j,t) \in [0,1]$

$zub(j,t) \in [0,1]$

$zlb(j,t) \leq zub(j,t)$

For a given stage of disease development j, the interval of expected test results may be in the normal range. In this case,

$$zlb(j,t) = zub(j,t) = 0 \tag{42}$$

Otherwise, the upper limit of test result t, zub(j,t) is greater than 0, creating a range for test values where

$$zlb(j,t) < zub(j,t) \tag{43}$$

Whereas the matrices H,A,S,Z represent the patient's information, the matrices H(i), A(j), S(j) and Z(j) represent the physician's knowledge of similar information relevant to disease i or disease stage j.

The Decision Model

The general objective of cluster analysis is to group similar data or variables into clusters. This is a sequential process, involving pairwise grouping of variables or clusters of variables. This clustering can continue until all data or variables are contained in the same overall cluster or can be designated to stop at a certain number of clusters. The cluster analysis used for this modelling is very controlled in that only a single grouping, involving a certain variable, is needed. This variable, a diseased patient, must be clustered with a specific disease at a given stage.

Applications of cluster analysis center on the formulation of a similarity matrix, Y, for the variables under consideration. In our research, these variables are the patient's sickness and the disease stage j,j = 1,2,...,g. The similarity matrix Y is a lower triangular matrix, whose structure is depicted in equation (44).

$$Y = \begin{bmatrix} y_{21} & & & \\ y_{31} & y_{32} & & \\ y_{41} & y_{42} & y_{43} & \\ \vdots & & & \ddots \\ y_{n1} & y_{n2} & \cdots & y_{n(n-1)} \end{bmatrix} \tag{44}$$

where

y_{ij} = similarity measure between disease stage i and disease stage j

n = total number of disease stages for all possible diseases plus 1.

The entities i and j of y_{ij}, where $1 \leq i \leq n - 1$, $1 \leq j \leq n - 1$, correspond to stages of the possible diseases. The entities i and j, where $i = j = n$, correspond to the patient's sickness or disease. Thus, the nth row or the last row of similarity matrix Y designates the similarity between the patient's sickness and possible diseases at each of their stages. The most similar entry in this row corresponds to the disease stage the patient is most likely to have.

The similarity measure y_{nj} between the patient's sickness and the possible diseases must be selected. The two most appropriate mathematical techniques include representation of these similarities using Minkowski metric distances and product moment correlation coefficients. Both of these are common similarity measures in cluster analysis theory. Of these, the Minkowski distance representation was

selected as the most appropriate technique for this research. This metric measurement permits the employment of the information contained in the matrices discussed in the previous sections. The Minkowski metric can be easily modified to incorporate symptom, sign, and test result severity bounds, as well as attribute - disease weighting. On the other hand, the correlation coefficient technique cannot handle the information of attributes possessing bounded severity intervals. Also, the disease-attribute weighting is not as intuitive in this latter method. Thus, the use of correlation coefficients as similarity measures has been deemed an alternate secondary approach to this problem.

The general Minkowski metric distance between a $1 \times m$ patient information matrix X_n and information matrix $X(j)$ of disease or disease stage j is

$$D_p(X(j), X_n) = \left[\sum_{k=1}^{m} |x(j,k) - x_n(k)|^p \right]^{1/p} \qquad (45)$$

where

$p \geq 1$,

m = number of attributes under consideration,

$x_n(k)$ = representation of attribute k in the patient,

$x(j,k)$ = representation of attribute k for disease or disease stage j,

$x_n(k) \in [0,1]$,

and $x(j,k) \in [0,1]$

Since $x_n(k)$ represents attribute k of the patient's sickness, it always takes on a single value. The variable x(j,k) may represent an aspect of history, a symptom, a sign, or a test result of disease or disease stage j. If this variable is dichotomous, then the formulation of equation (45) is adequate to handle all possible cases of attribute absence or presence. But often x(j,k) represents nonbinary attributes. The value associated with this variable might then be an interval of severity rather than a single measure. For these instances, let [xlb(j,k), xub(j,k)] represent the severity interval of variable x(j,k).

Consider the following cases where x(j,k) is represented as a bounded interval of severity. If

$$xlb(j,k) \leq x_n(k) \leq xub(j,k) \quad , \tag{46}$$

then the patient's attribute is in the interval of severity for disease stage j. For this case,

$$x(j,k) - x_n(k) = xlb(j,k) - x_n(k) = xub(j,k) - x_n(k) = 0 \ . \tag{47}$$

Suppose the patient's attribute k lies outside the interval of severity for disease stage j. If $x_n(k) < xlb(j,k)$ then $x(j,k) - x_n(k)$ of equation (45) becomes $xlb(j,k) - x_n(k)$. If $x_n(k) > xub(j,k)$, then $x(j,k) - x_n(k)$ of equation (45) becomes $xub(j,k) - x_n(k)$. The following sets can be constructed to assist in representing the dichotomous and non-dichotomous attribute comparisons for disease stage j. Let

$$C_j = \{k | x(j,k) \in \{0,1\},\ x_n(k) \in \{0,1\},\ 1 \le k \le m\} \tag{48}$$

$$F_j = \{k | x_n(k) < x\ell b(j,k),\ 1 \le k \le m\}$$

$$G_j = \{k | x_n(k) > xub(j,k),\ 1 \le k \le m\}$$

Equation (45) can now be written as

$$D_p(X(j),\ X_n) = \Big[\sum_{k \in C_j} |x(j,k) - x_n(k)|^p + \sum_{k \in F_j} |xlb(j,k) - x_n(k)|^p + \sum_{k \in G_j} |xub(j,k) - x_n(k)|^p \Big]^{1/p} \tag{49}$$

In this diagnosis decision process, attribute k of disease or disease stage j has a specific weight of importance to the diagnosis of that disease or disease stage. This fuzzy concept was termed the "pertinence" of attribute k to the diagnosis of disease or disease stage j in Chapter II. Let Ψ be the set of attributes under consideration for the diagnoses. Construct a fuzzy set, ξ_j, where ξ_j equals the set of "pertinent" attributes of disease j. The membership function of this fuzzy set determines the weight associated with attribute k for disease or disease stage j. Thus $u_{\xi_j}(k) = w(j,k)$ where w(j,k) is an entry in a n x m matrix W of weights. Thus

$$\Psi \xrightarrow{u_{\xi_j}(\Psi)} [0,1] \quad j = 1,2,\ldots,n \ . \tag{50}$$

By choosing various values of p, different metric distance functions can be obtained from the Minkowski metric equation. In this study, the familiar Euclidean distance measure is formed by setting $p = 2$.

By incorporating weights and the Euclidean distance assumption into equation (49), the revised Minkowski metric becomes

$$D_2(X(j), X_n) = \Big[\sum_{k \in C_j} |w(j,k)\,(x(j,k) - x_n(k))|^2 + \sum_{k \in F_j} |w(j,k)\,(xlb(j,k) - x_n(k))|^2 + \sum_{k \in G_j} |w(j,k)\,(xub(j,k) - x_n(k))|^2 \Big]^{1/2} \qquad (51)$$

where

$$w(j,k) \in [0,1]$$
$$xub(j,k) \in [0,1]$$
$$xlb(j,k) \in [0,1]$$
$$x_n(k) \in [0,1]$$
$$x(j,k) \in \{0,1\}$$

This measure of similarity

$$D_2(X(j), X_n) = y_{jn} \qquad (52)$$

of the similarity matrix Y. The minimum revised Minkowski metric for the disease stages under consideration reflects the most similar

disease - patient pair. The objective of the diagnosis decision is to

$$\text{minimize } D_2(X(j),X_n),\ j = 1,2,3,\ldots,n \tag{53}$$

Equation (51) is thus the general diagnosis decision model for this study.

The matrices presented earlier in this chapter are easily incorporated into this general diagnosis equation to determine the various diagnoses. For a decision upon a medical hypothesis,

$$D_2\ (X(j),\ X_n) = D_2\ (H(i),\ H) + D_2(A(j),A) \tag{54}$$

where

$j \in$ {possible disease stages}

$i \in$ {possible diseases}

For the initial preliminary diagnoses,

$$D_2(X(j),X_n) = D_2(H(i),H) + D_2(A(j),A) + D_2(S(j),S) \tag{55}$$

The metric distance structure for other preliminary and final diagnoses is

$$D_2(X(j),X_n) = D_2(H(i),H) + D_2(A(j),A) + D_2(S(j),S) + D_2(Z(j),Z) \tag{56}$$

Thus, these various diagnoses are simple to determine due to the organization of the data structure.

By minimizing the function $D_2(X(j),X_n)$, the best candidate for the diagnosis is found. As was previously stated, the earlier diagnoses

may limit the diseases to a group of possibilities or just eliminate certain candidates. Therefore, the k smallest entries in the last row of the similarity matrix Y gives the set P of the k most likely diseases to be present. Another possible technique for determining a set P of the most likely diseases is to set a threshold value,h, such that

$$P = \{j \mid D_2(X(j),X_n) \leq h,\ h > 0\} \tag{57}$$

The desired strategy must be selected for the situation under consideration.

This chapter has developed abstract mathematical formulations of the various aspects of the diagnosis decision process. Illustrations and examples of this modelling will be presented in Chapters IV and V.

CHAPTER IV

MODELLING A DIAGNOSIS DECISION PROCESS: AN INSTRUCTIVE EXAMPLE

The model development of Chapter III would be somewhat incomplete if an instructive example of the information structures and the diagnosis model were not given. Due to the vast amount of medical knowledge and the data required for a model of the complete process, the example herein is concerned only with reaching a medical hypothesis as to which disease is present in the patient. The patient information needed for this example includes the patient's past history and present symptoms of his sickness. The diseases under consideration for this medical hypothesis have been limited to rheumatic valvular heart diseases. Even for this very specialized area of medicine, the possible information structures needed to reach a final diagnosis would be immense, making an earlier diagnosis decision the only feasible one for this research.

The information and data used for this example were obtained from prominent medical references and were evaluated and supplemented as appropriate by local physicians. Since large amounts of actual patient information were not available for this research, the data presented often tends to be subjective in nature. As a result, the following exemplification is representative of the actual process from the viewpoint of a limited number of physicians used as well as pertinent references from medical literature.

Example of a Hypothesis Decision Process

General

In the example given here, the patient's sickness may be clustered with one of five possible diseases at one of three disease development stages. Therefore there are 15 stages of diseases or disease variables that can be paired with the patient's sickness. The five rheumatic valvular diseases are as follows:

1. Aortic insufficiency
2. Aortic stenosis
3. Mitral insufficiency
4. Mitral stenosis
5. Tricuspid insufficiency

All of these diseases are related in that they are concerned with malfunctions around the valve areas of the heart, frequently causing heart failure. Each disease is assumed to have three stages of development - early, intermediate, and late. Although these are not naturally disjoint categories, medical reasons involving the size of the valve openings warrant such divisions.

Relevant Past History

There are relatively few aspects of patient's past history which are relevant to the diagnoses of rheumatic valvular heart diseases. These aspects, which make up the set Ω in Chapter III, are listed in Table 1. The nonmedical aspects, including sex and age, are important to many diagnoses and thus should always be given consideration. The medical aspects are made up of diseases which directly cause valvular heart disease. These include rheumatic fever and syphilis. Other

causes of these diseases, such as atherosclerosis, exist, but do not show up in the patient's past history.

Often the patient might have had rheumatic fever or syphilis, but it went unobserved by a physician and thus was not medically diagnosed and designated. To alleviate this, the fuzzy sets θ_j are created and contain "prominent" symptoms of rheumatic fever and syphilis. These sets are shown in Table 2. Note that each of these sets only contains two elements. The binary matrix V(j), representing the presence or absence of the elements of θ_j, must be mapped into the patient history matrix H. As in equation (7), this is symbolically written as

$$V(j) \xrightarrow{f(V(j))} h(j) \tag{58}$$

where

$$f(V(j)) \in \{0,1\}$$

In this example V(j), for j = 1, 2, are only 1 x 2 matrices, allowing the function f to be very simplistic in nature. If both the symptoms of θ_j are present in the patient's past history, then the disease j is assumed present and h(j) = 1. Thus

$$f(V(j)) = \begin{cases} 1 \text{ if } V(j) = [1,1] \\ 0, \text{ otherwise} \end{cases} \tag{59}$$

The past history of the patient can then be incorporated into a 1 x 8 binary matrix, H, which denotes the presence or absence of the history aspects in Table 1. This history matrix must be compared to the binary history matrix H(i), for each of the five diseases,

Table 1. Relevant Aspects of the Patient's Past History.

Medical	Nonmedical
Diseases	Sex
1. Rheumatic fever	4. Male
2. Syphilis	5. Female
3. None of these diseases	Age
	6. Young (Under 30)
	7. Middle Age (30-60)
	8. Old (Over 60)

Table 2. Prominent Symptoms for Past Undiagnosed Diseases.

Rheumatic Fever (j=1)	Syphilis (j=2)
1. High fever	3. Pinkish skin rash
2. Sore, sensitive joints	4. Pronounced sores

$i = 1,2,\ldots,5$. Past history is dependent upon the disease and not the stage of the disease, so 15 matrices are not needed. The history matrices for each of the diseases are given in Table 3. They are very similar to one another, since the disease area under consideration is restricted. The medical aspects of the matrix take on a unit value if the past disease relates to the valvular disease, a zero value if not. The nonmedical aspects for these diseases are all present to some degree in past cases of valvular heart diseases, giving all of them a unit value. The similarities between these matrices and the patient's matrix must be determined. The weights, $w(i,k)$, associated with the importance of history aspect k to disease i are given in Table 4. These weights reflect the percentage of time the history aspect k has been present in a patient with disease i [2]. If history aspect k does not influence disease i, i.e. $h(i,k) = 0$, then $w(i,k)$ takes on a negative unit value. By incorporating these weights into equation (51), the similarity measure $D_2(H(i), H)$, $i = 1,2,\ldots,5$, can be determined. Thus

$$D_2(H(i),H) = \left[\sum_{k=1}^{8} |w(i,k)\ (h(i,k) - h(k))|^2 \right]^{1/2}, \qquad (60)$$

$$i = 1,2,\ldots,5$$

These values are used later to determine the hypothesis decisions.

Present Symptoms of the Diseases

The other collection of information needed to reach a medical hypothesis pertains to patient related problems that are transformed into medically designated symptoms of valvular heart disease. The set of patient problem areas Π, relevant to these diseases, is given in Table 5.

Table 3. Binary Matrices of Past Patient History.

Disease i	Disease History Vector-H(i)							
k =	1	2	3	4	5	6	7	8
1. Aortic Insufficiency	1	1	1	1	1	1	1	1
2. Aortic Stenosis	1	0	1	1	1	1	1	1
3. Mitral Insufficiency	1	0	1	1	1	1	1	1
4. Mitral Stenosis	1	0	1	1	1	1	1	1
5. Tricuspid Insufficiency	1	0	1	1	1	1	1	1

Note: k denotes history aspect number from Table 1.

Table 4. Weights of History Aspects for Each Disease.

Disease i	Weights-w(i,k)							
k =	1	2	3	4	5	6	7	8
1. Aortic Insufficiency	.6	.2	.2	.7	.3	.2	.7	.1
2. Aortic Stenosis	.7	-1	.3	.7	.3	.2	.7	.1
3. Mitral Insufficiency	.6	-1	.4	.6	.4	.3	.6	.1
4. Mitral Stenosis	.6	-1	.4	.2	.8	.3	.6	.1
5. Tricuspid Insufficiency	.6	-1	.4	.4	.6	.3	.6	.1

For each of these problem areas, there is one or more corresponding medical symptoms. The designated symptoms of rheumatic valvular heart disease, β, are given in Table 6. The definitions of these symptoms, as well as factors influencing their severity, are given in Appendix B. Two purposes for the previous mathematical model were to map the problems of Table 5 into the symptoms of Table 6 and to determine the severity levels of these symptoms. The information needed to achieve this must be obtained from the patient and incorporated into a profile, Δ_q, for each problem q. The profile categories needed for this example are as follows:

1. Location of the problem,
2. Continuity of the problem,
3. Activities accompanying symptom occurrence,

and 4. Specifics for severity determination (fuzzy sets).

Table 7 gives a complete listing of the specific factors for each of these categories. Therefore, for each problem communicated by the patient, the listed factors assist in determining the symptom and the symptom's level of severity. Numbers associated with each of these factors, found in Table 7, are placed in the appropriate factor rows and problem columns of the matrix B, defined in equation (12).

For this exemplification, the profile matrix B is a 10 x 15 matrix.

$$B = [b(i,j)], \quad \begin{matrix} i = 1,2,\ldots,10 \\ j = 1,2,\ldots,15 \end{matrix} \tag{61}$$

Table 5. Patient Related Problem Areas.

1. Breathing difficulty; Shortness of breath	9. Dizziness; Lightheadedness; Confusion; Greying out; Unconsciousness
2. Coughing	10. Irregular heartbeat
3. Blood in sputum	11. Pain
4. Purple skin color	12. Swelling
5. Red face and lips	13. Weight gain
6. Yellow skin color	14. Appetite loss
7. Weakness; Lack of energy	15. Nausea and Vomiting
8. Fatigue; Tiredness	

Table 6. Medically Designated Symptom.

1. Exertional Dyspnea	11. Palpitations
2. Paroxysmal Nocturnal Dyspnea	12. Typical Anginal Pain
3. Cough	13. Epigastric Anginal Pain
4. Hemoptsis	14. Pain from Enlarged Liver
5. Cyanosis	15. Headaches
6. Malar Flush	16. Peripheral Edema
7. Jaundice	17. Ascites
8. Weakness	18. Weight Gain from Edema
9. Fatigue	19. Anorexia
10. Syncope	20. Nausea and Vomiting

Table 7. Factors of Problem and Symptom Profiles.

Location Factors	
1. Fingernails	9. Lower chest
2. Face	10. Middle chest
3. Lips	11. Inner left arm
4. Ankles	12. Arms
5. Feet	13. Back of Neck
6. Legs	14. Head
7. Abdomen	15. Back
8. Upper abdomen	

Continuity Factors	
1. Continuous	2. Intermittent

Activities Provoking or Accompanying Intermittent Symptom Occurrence	
1. Running	6. Resting
2. Climbing stairs	7. Recumbency
3. Lifting heavy weight	8. Sleep
4. Walking	
5. Standing from sitting position	

Specifics for Severity Determination	
1. Productivity of cough	5. Purpleness of Skin
2. Painfulness of Headache	6. Awareness of patient
3. Bloodiness of Sputum	
4. Increasing weight	

where in (61),

$$b(i,j) = \begin{cases} 0 \text{ if problem } j \text{ is present} \\ 1 \text{ if problem } j \text{ is absent} \end{cases}$$

$b(2,j) \in [0,1]$ (Severity specifics for problem j-fuzzy sets)

$$b(3,j) = \begin{cases} 1 \text{ if problem } j \text{ is continuous} \\ 2 \text{ if problem } j \text{ is intermittent} \end{cases}$$

$$\left.\begin{matrix} b(4,j) \\ b(5,j) \end{matrix}\right\} \in \{\text{Activities accompanying problem } j \text{ occurrence}\}$$

$$\left.\begin{matrix} b(6,j) \\ \vdots \\ b(10,j) \end{matrix}\right\} \in \{\text{Locations of problem } j\}$$

Before determining the level of a symptom, the symptom itself must be designated based on the problem and problem profile. For valvular heart disease, the greatest amount of information needed to do this includes the problem, the locations of the problem, and the activity accompanying the problem occurrence. In Tables 8 and 9, problems are associated with their designated symptoms. Many problems and symptoms are in one to one correspondence with one another and need no additional information from the problem profile. But for other problems, especially pain, additional factors directly influence the designation of the symptom. In these instances, the problems are mapped to the symptoms via determination of the factor combinations in the patients. It should be noted that there are many other factors

Table 8. Designation of Symptoms from Problems and Problem Profiles.

Problems	Symptoms	Factors of Problem		Severity Variations
		Location*	Activity*	
1. Breathing difficulty; Shortness of breath	1. Exertional dyspnea	0	1-7	Yes
	2. Paroxysmal nocturnal dyspnea	0	8	No
2. Cough	3. Cough	0	0	Yes
3. Blood in sputum	4. Hemoptsis	0	0	Yes
4. Purple skin color	5. Cyanosis	0	0	Yes
5. Redface and lips	6. Malar flush	0	0	No
6. Yellow skin color	7. Jaundice	0	0	No
7. Weakness; Lack of energy	8. Weakness	0	0	Yes
8. Fatigue; Tiredness	9. Fatigue	0	0	Yes
9. Dizziness; Lightheadedness; Confusion; Greying out; Unconsciousness	10. Syncope	0	0	Yes
10. Irregular heartbeat	11. Palpitations	0	0	Yes
11. Pain	12. Typical Anginal Pain	10 11	1-7	Yes
	13. Epigastric Anginal Pain	7 15	1-5	No
	14. Pain from Enlarged Liver	8 9	0	No
	15. Headaches	13 14	0	Yes

*Note that the factors are represented by their numbers from Table 7.

Table 9. Designation of Symptoms from Problems and Problem Profiles.

Problems	Symptoms	Factors of Problem		Severity Variations
		Location*	Activity*	
12. Swelling	16. Peripheral edema	4,5,2	0	Yes
	17. Ascites	7	0	No
13. Weight gain	18. Weight gain from edema	0	0	Yes
14. Appetite loss	19. Anorexia	0	0	No
15. Nausea and Vomiting	20. Nausea and Vomiting	0	0	No

*Note that the factors are represented by their numbers from Table 7.

associated with each of these problems, but only those listed aid in discriminating between symptoms.

Once the symptom has been designated, the severity of the symptom must be determined. Granted, if a symptom k has not been designated, then its level of severity, a(k), equals zero. If the symptom is dichotomous and its severity representation binary, then a designated symptom k would have a level of severity a(k) equal to one. Such dichotomous symptoms for this example are revealed in the last column of Tables 8 and 9, where severity variations are assumed not to exist. Determination of the levels of severity for the remaining non-dichotomous symptoms is more complex, but is aided through the use of fuzzy set theory.

Various fuzzy sets have been evaluated to aid in determining the severities of nonbinary symptoms in this example. The membership functions for each of these fuzzy sets directly or indirectly determine one or more symptoms' severity. Tables 10 and 11 illustrate these fuzzy sets and their membership functions. For each fuzzy set, a finite set of representative elements, x_i, $x_i \in X$, have been selected for illustration. The membership functions have been quantified using a distance criterion similar to that developed by Kochen [58,59] to quantify fuzzy adjectives. To illustrate this technique, consider the first fuzzy set, B1, denoting the strenuousness of an activity. At one end of a line, a stationary point denoting recumbency (no activity) is designated. A mark is then placed at the point on the line where the distance between points quantifies the strenuousness of the activity x_i with respect to recumbency. When compared with other activities, the distance between

Table 10. Fuzzy Sets Used to Determine Severities of Symptoms.

Nonfuzzy Set-X	Fuzzy Set-B	Symbol of Fuzzy Set	x_i	$u_B(x_i)$
Activities	Strenuous Activities	B1	Running	1.0
			Climbing stairs	.93
			Lifting heavy weight	.96
			Walking	.36
			Standing from sitting position	.14
			Resting	.12
			Recumbency	0.0
Pain	Painfulness	B2	Mild Pain	.17
			Pain	.5
			Severe Pain	1.0
			No Pain	0.0
Blood	Bloodiness	B3	A little blood	.23
			Flects of blood	.39
			Slight bleeding	.23
			Large amounts of blood	1.0
			Profuse bleeding	1.0
			Blood	.54
			No Blood	0.0

Table 11. Fuzzy Sets Used to Determine Severities of Symptoms.

Nonfuzzy Set-X	Fuzzy Set-B	Symbol of Fuzzy Set	x_i	$u_B(x_i)$
Weight	Increasing Weight	B4	5 lbs.	.28
			10 lbs.	.5
			15 lbs.	.75
			20 lbs.	1.0
			A little weight	.22
			A lot of weight	.55
			No weight loss	0.0
Colors	Purpleness	B5	Ashen	.09
			Blue tinge	.32
			Purple tinge	.51
			Purple	.77
			Very blue grey	.21
			Very purple	1.0
			No Purple	0.0
States of Patient	Awareness of Patient	B6	Unconsciousness	0.0
			Confusion	.61
			Greying out	.41
			Dizziness	.78
			Lightheadedness	.82
			Normal awareness	1.0

recumbency and x_i represents the relative strenuousness of the activity. Since a distance criterion is used in the similarity matrix for the diagnosis decision, consistency is maintained by using a distance measure for membership functions of a fuzzy set.

Once these fuzzy sets have been evaluated, their membership functions should be used to determine symptom severity. The severity of many symptoms caused by rheumatic valvular diseases increases in accordance with the membership function. Such fuzzy sets include bloodiness of the sputum, purpleness of the skin, pain of the headache and others. The severity of other symptoms decreases with the increasing membership function. Examples of these fuzzy sets include strenuousness of the activity and awareness of the patient. Many symptoms under consideration have severities which reflect the minimum amount of activity which induces the symptom. By taking the complement of the fuzzy set of strenuous activities, the proper severity function results. The severity levels of syncope are also determined via this complement technique with regard to the patient's awareness. Table 12 better illustrates the use of these fuzzy sets and the functions to denote symptom levels, a(k).

As can be seen from this table, most of the symptom severities are influenced by fuzzy sets. Many depend only upon the fuzzy sets while others are simple functions of fuzzy sets and location factors. The severity levels of coughs and peripheral edema have been effectively determined without the use of fuzzy set theory. In these cases, nonfuzzy factors are sufficient enough to quantify the severity levels. Again, note that many of these symptoms possess other factors, but only

Table 12. Severity Determination of Symptoms.

Symptom i*	Factors Influencing Severity**					Severity Level-a(i)
	Location	Continuity	Activity	Severity Specifics	Fuzzy Sets	
1. Exertional Dyspnea	0	0	1-7	0	X	$a(1) = 1 - u_{B1}(x_i)$
3. Cough	0	2	1-4	0		$a(3) = .14$
	0	2	7	0		$a(3) = .62$
	0	2	1-4	1		$a(3) = .79$
	0	2	7	1		$a(3) = .89$
	0	1	0	1		$a(3) = 1.0$
4. Hemoptsis	0	0	0	3	X	$a(4) = u_{B3}(x_i)$
5. Cyanosis	1,2,3 5,6,12	0	0	6	X	$a(5) = u_{B5}(x_i)/2$
	7,10	0	0	6	X	$a(5) = (u_{B5}(x_i)/2) + .5$
8. Weakness	0	0	1-6	0	X	$a(8) = 1 - u_{B1}(x_i)$
9. Fatigue	0	0	1-5	0	X	$a(9) = 1 - u_{B1}(x_i)$
10. Syncope	0	0	0	7	X	$a(10) = 1 - u_{B6}(x_i)$
11. Palpitations	0	0	1-7	0	X	$a(11) = 1 - u_{B1}(x_i)$
12. Typical Anginal Pain	10	0	1-7	0	X	$a(12) = (1 - u_{B1}(x_i))/2$
	10,11	0	1-7	0	X	$a(12) = (1 - u_{B1}(x_i))/2 + .5$

Table 12. Severity Determination of Symptoms.
(Continuation)

Symptom i*	Factors Influencing Severity**					Severity Level-a(i)
	Location	Continuity	Activity	Severity Specifics	Fuzzy Sets	
15. Headaches	0	0	0	2	X	$a(15) = u_{B2}(x_i)$
16. Peripheral Edema	4,5	2	0	0		$a(16) = .24$
	4,5	1	0	0		$a(16) = .41$
	2	2	0	0		$a(16) = .79$
	2	1	0	0		$a(16) = 1.0$
18. Weight Gain from Edema	0	0	0	4	X	$a(18) = u_{B4}(x_i)$

*Note that these numbers correspond to numbers from Tables 8 and 9.

**Note that these factors are given as numbers, corresponding to those in Table 7.

those pertinent to severity determination are given.

Once the severity level, a(k), for each symptom k has been measured, the patient's symptom matrix, A, must be compared to the symptom matrix A(j), for disease stage j. In Tables 13 through 16, the lower bound, alb(j,k) and upper bound aub(j,k) of each symptom k for each disease stage j are given, where the diseases stages are numbered consecutively. The severity bounds for nonbinary symptoms are based on the same measures used to determine the patient's symptom severities. Weights of importance or pertinence, w(j,k), for each symptom k of disease stage j are given in Table 17. Although these weights were subjectively determined, they greatly aid in stressing the importance of the symptom's presence or absence from the patient. The positive weights of this table denote the importance of the symptoms being present for the given disease candidate. Conversely, the negative weights denote the importance of the symptom not being present for the disease candidate. Using these severity bounds and weights, the similarities between the patient's symptoms and the symptoms of disease stage j are determined by $D_2(A(j),A)$, $j = 1,2,\ldots,15$. Thus, from equation (51),

$$D_2(A(j),A) = \Bigg[\sum_{k \in C_j} |w(j,k)\ (alb(j,k) - a(k))|^2 + \sum_{k \in F_j} |w(j,k)\ (alb(j,k) - a(k))|^2 + \sum_{k \in G_j} |w(j,k)\ (aub(j,k) - a(k))|^2 \Bigg]^{1/2} \qquad (62)$$

Table 13. Symptom Matrices for the Stages of Valvular Heart Disease.

Disease Stage-j	alb (j,1)	aub (j,1)	alb (j,2)	aub (j,2)	alb (j,3)	aub (j,3)	alb (j,4)	aub (j,4)	alb (j,5)	aub (j,5)
1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.7	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.6	1.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.2	0.0	0.0	0.7	0.9	0.0	0.3	0.0	0.3
6	0.0	1.0	0.0	0.0	0.9	1.0	0.3	1.0	0.2	0.4
7	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.6	0.0	0.0	0.7	0.9	0.0	0.0	0.0	0.0
9	0.6	1.0	1.0	1.0	0.9	1.0	0.0	0.3	0.0	0.0
10	0.0	0.7	1.0	1.0	0.0	0.9	0.0	0.3	0.0	0.3
11	0.5	1.0	1.0	1.0	0.7	0.9	0.3	1.0	0.0	0.6
12	0.7	1.0	1.0	1.0	0.9	1.0	0.3	1.0	0.5	1.0
13	0.0	0.7	0.0	0.0	0.7	0.9	0.0	0.0	0.0	0.3
14	0.5	0.7	0.0	0.0	0.7	0.9	0.3	0.6	0.2	0.6
15	0.7	1.0	0.0	0.0	0.9	1.0	0.3	1.0	0.6	1.0

Table 14. Symptom Matrices for the Stages of Valvular Heart Disease.

Disease Stage-j	alb (j,6)	aub (j,6)	alb (j,7)	aub (j,7)	alb (j,8)	aub (j,8)	alb (j,9)	aub (j,9)	alb (j,10)	aub (j,10)
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.2
2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.7	0.1	0.4
3	0.0	0.0	0.0	0.0	0.0	0.1	0.5	0.7	0.1	0.4
4	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.6	0.1	0.4
5	0.0	0.0	0.0	0.0	0.5	0.9	0.6	1.0	0.4	1.0
6	0.0	0.0	0.0	0.0	0.8	1.0	0.8	1.0	0.5	1.0
7	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0
8	0.0	0.0	0.0	0.0	0.3	0.7	0.5	0.8	0.0	0.0
9	0.0	0.0	0.0	0.0	0.7	1.0	0.8	1.0	0.0	0.0
10	1.0	1.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0
11	1.0	1.0	1.0	1.0	0.0	0.1	0.0	0.1	0.0	0.0
12	1.0	1.0	1.0	1.0	0.1	1.0	0.5	1.0	0.0	0.0
13	0.0	0.0	1.0	1.0	0.6	0.9	0.6	1.0	0.0	0.0
14	0.0	0.0	1.0	1.0	0.6	0.9	0.8	1.0	0.0	0.0
15	0.0	0.0	1.0	1.0	0.8	1.0	0.8	1.0	0.0	0.0

Table 15. Symptom Matrices for the Stages of Valvular Heart Disease.

Disease Stage-j	alb (j,11)	aub (j,11)	alb (j,12)	aub (j,12)	alb (j,13)	aub (j,13)	alb (j,14)	aub (j,14)	alb (j,15)	aub (j,15)
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
3	0.6	1.0	0.0	1.0	1.0	1.0	1.0	1.0	0.4	1.0
4	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.7	0.3	1.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.1	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.7	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	0.6	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.6	1.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.2
14	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.2	0.6
15	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.6	1.0

Table 16. Symptom Matrices for the Stages of Valvular Heart Disease.

Disease Stage-j	alb (j,16)	aub (j,16)	alb (j,17)	aub (j,17)	alb (j,18)	aub (j,18)	alb (j,19)	aub (j,19)	alb (j,20)	aub (j,20)
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0
6	0.3	0.5	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.5	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
12	0.4	0.8	1.0	1.0	0.3	0.6	1.0	1.0	1.0	1.0
13	0.0	0.5	1.0	1.0	0.0	0.3	1.0	1.0	0.0	0.0
14	0.4	0.8	1.0	1.0	0.3	0.8	1.0	1.0	1.0	1.0
15	0.7	1.0	1.0	1.0	0.5	1.0	1.0	1.0	1.0	1.0

Table 17. Pertinent Weights of Each Symptom to the Diagnosis of the Disease Stages.

Symptom k	Disease Stage-j 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.4	.8	1.0	-.1	.6	1.0	.6	.8	1.0	.8	1.0	1.0	.6	.6	.6
2	-.1	.4	.8	-.1	-.1	-.1	-.1	-.1	.7	.3	.6	.9	-.1	-.1	-.1
3	-.8	-.6	-.4	-.1	.6	.8	-.4	.4	.6	.8	.8	1.0	.3	.6	.8
4	-.8	-.6	-.3	-.6	.2	.8	-.6	-.4	.2	.5	.8	1.0	.1	.7	1.0
5	-.8	-.4	-.1	-.1	.5	.7	-.8	-.4	-.1	.4	.7	.9	.6	.7	.9
6	-.8	-.6	-.4	-.8	-.4	-.1	-.5	-.3	-.1	.3	.5	.7	-.1	-.1	-.1
7	-.8	-.8	-.2	-.8	-.8	-.2	-.6	-.6	-.2	.3	.5	.7	.5	.7	.9
8	-.1	.3	.8	.5	.8	1.0	.3	.5	.7	.4	.7	.9	.5	.8	.8
9	.3	.5	.8	.5	.8	1.0	.3	.5	.7	.4	.7	.9	.5	.8	.8
10	.5	.6	.9	.5	.7	.9	-.6	-.6	-.4	-.6	-.6	-.4	-.1	-.1	-.1
11	-.4	.5	.7	-.6	-.4	.6	.3	.5	.8	.5	.7	1.0	-.1	-.1	-.1
12	-.4	-.1	.6	.4	.6	.9	-.8	-.8	-.6	-.5	-.3	-.1	-.1	-.1	-.1
13	-.1	-.1	.7	-.1	-.1	-.1	-.8	-.8	-.6	-.5	-.3	-.1	-.1	-.1	-.1
14	-.8	-.6	-.4	-.8	-.6	-.2	-.8	-.6	-.2	-.4	-.1	.6	.6	.8	1.0
15	-.2	.4	.8	-.4	-.4	-.2	-.4	-.4	-.4	-.4	-.4	-.1	.4	.6	.8
16	-.8	-.4	-.1	-.1	.4	.8	-.6	-.3	-.1	-.1	.4	.8	.6	.8	1.0
17	-.8	-.4	-.2	-.8	-.4	-.2	-.8	-.4	-.2	-.6	-.1	.6	.6	.8	1.0
18	-.6	-.4	-.1	-.1	.4	.8	-.6	-.3	-.1	-.1	-.4	.6	.6	.8	1.0
19	-.6	-.3	-.1	-.3	-.2	-.1	-.6	-.3	-.1	.3	.4	.7	.4	.6	.8
20	-.8	-.6	-.3	-.7	-.3	-.3	-.7	-.5	-.3	-.1	.4	.7	.4	.6	.8

$$j = 1,2,\ldots,15$$

where

$$C_j = \{k \mid alb(j,k) = aub(j,k) \in \{0,1\}\ ,$$

$$a(k) \in \{0,1\},\ 1 \leq k \leq 20\}$$

$$F_j = \{k \mid a(k) < alb(j,k),\ 1 \leq k \leq 20\}$$

$$G_j = \{k \mid a(k) > aub(j,k),\ 1 \leq k \leq 20\}$$

$$w(j,k) \in [0,1]$$

$$aub(j,k),\ alb(j,k),\ a(k) \in [0,1]$$

By combining the similarity measures of patient history and symptoms, the similarity measure y_{jn}, needed to reach a medical hypothesis, can be determined. Consequently,

$$y_{jn} = \{D_2(H(i),H) + D_2(A(j),A)\} \tag{63}$$

$$i = 1,2,\ldots,5$$

$$j = 1,2,\ldots,15$$

The solution to the following problem yields the most likely disease stage candidate for the medical hypothesis.

$$\text{Min } y_{jn} \tag{64}$$

$$j = 1,2,\ldots,15$$

Other strategies, like that represented in equation (57), yield a group of the most possible disease stages, rather than a single disease stage.

This chapter has presented an illustrative example of the diagnosis decision process and its modelling representations. To determine the usefulness of this developed model, diagnoses using information derived from this example are analyzed in Chapter V.

CHAPTER V

COMPUTERIZATION AND VALIDATION OF THE FUZZY DIAGNOSIS MODEL

By incorporating the information of the illustrative example into the diagnosis decision model, hypothesis decisions involving valvular heart diseases can be reached. This chapter attempts to validate the diagnosis model by analyzing such diagnosis decisions. The purpose of this validation is to obtain some insight into the model's performance and to identify any problem areas or deficiencies of the model. To make such a validation feasible, computerization of the fuzzy diagnosis model is considered necessary. Consequently, a computer program for the diagnosis of valvular heart diseases via fuzzy set theory is offered herein. By contrasting the diagnosis of the computerized model with the diagnoses of a physician, the validity or preciseness of the proposed model can be studied.

Computerization of the Model

The computer program developed for this phase of the study centers upon the mathematical modelling and medical data presented in Chapter IV. The flowchart used to formulate this program is illustrated in Figure 2. A complete listing of the Fortran programming used to reach hypothesis decisions may be found in Appendix C.

The computerized program is itself interactive in nature, taking on some of the sequential aspects of the actual decision process.

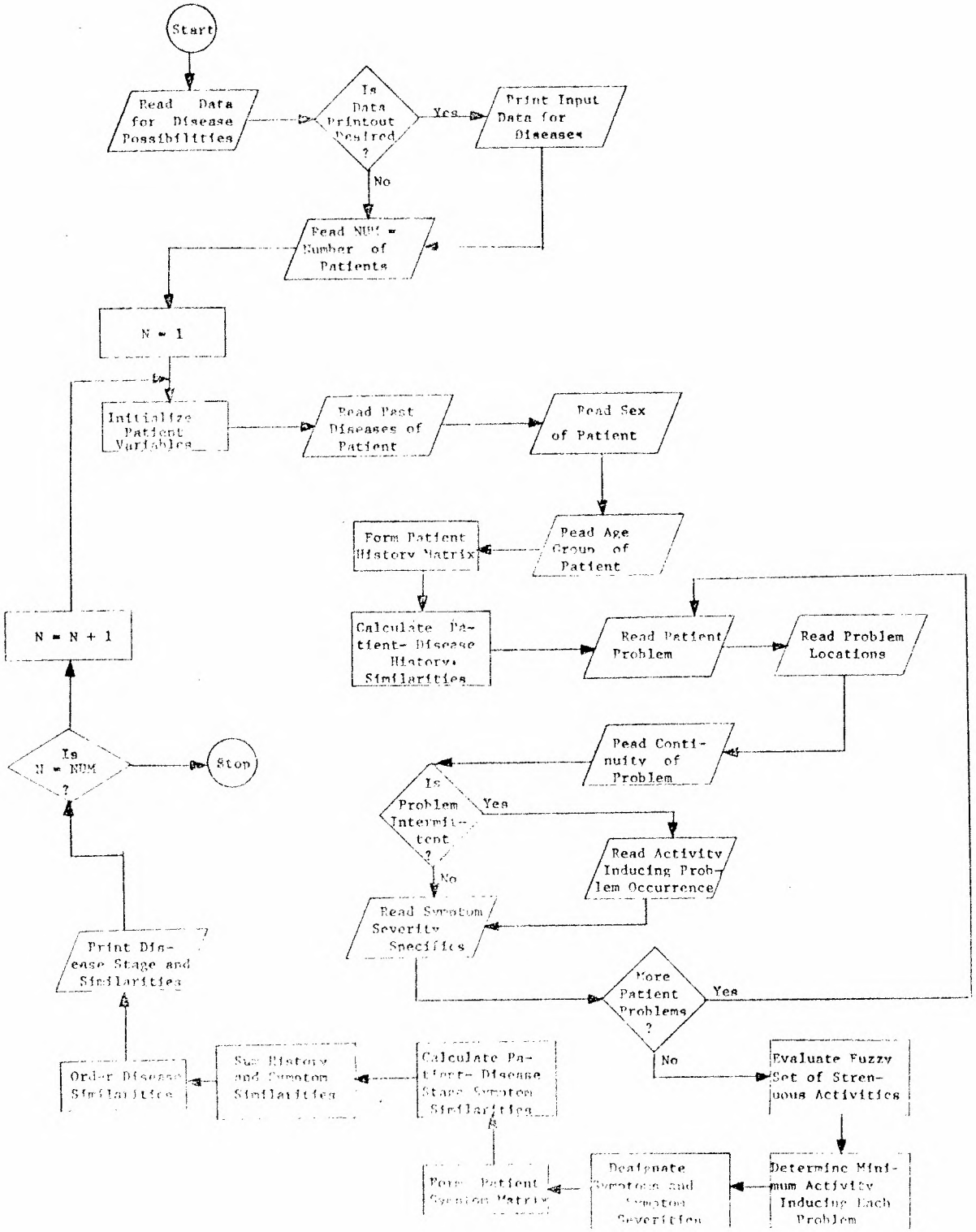

Figure 2. Flowchart of the Computerized Diagnosis Decision Model.

Pertinent questions concerning past medical and nonmedical history, as well as symptom aspects are asked. Patient problems, problem factors, and fuzzy severity aspects are put forth singularly, so that full information concerning each specific problem may be discerned. In this way, all important information for symptom and symptom severity determination is obtained. Answers to each of these questions must be derived from answer categories supplied by the computer output or from specific answer sheets supplied with the program in Appendix C. Membership functions of fuzzy sets, problem location factors, continuity determinators, problem inducing factors, and problem severity specifics must be read from these sources or listings and used as input data at the particular point of the interaction.

The various output of this program directly aids the physician in reaching a hypothesis. Such program output includes:

1. Measured patient history matrix, H.
2. Measured patient symptom matrix, A.
3. Ordered table of patient - disease stage similarity measures,

$$D(X(j),X_n)$$

where

$$D(X(j),X_n) = D(H(i),H) + D(A(j),A) \qquad (65)$$

$$i = 1,2,\ldots,5$$

$$j = 1,2,\ldots,15.$$

Quantified patient matrices should give the physician a better idea of the information present in the decision, since the matrices are in such a concise form. By having a table of ordered disease candidates,

the physician should be better able to determine his future strategy and possible alternate disease candidates. Singular hypothesis candidates or multiple candidates may result from close evaluation of this table. A summary of this computerized model's input and output is given in Figure 3.

Model Validation

As was previously mentioned, the fuzzy diagnosis model is validated by comparing computerized hypotheses and physician-made hypotheses when the same information is used for both decisions. Initially, patient medical files were sought as the source for physician made decisions. History and symptoms could have been extracted from these files and used as input for the computerized decision. Unfortunately, due to an increased protection and confidentiality of patient information in the local hospitals, such patient files were unobtainable. Aside from this fact, valvular heart disease does not occur prolifically, thus making it difficult to obtain vast quantities of information. Therefore, some other appropriate and effective technique had to be devised to validate this model.

Consequently, this fuzzy diagnosis decision model has been validated via comparisons with mock physician diagnosis decisions. This process involves selecting a group of history aspects and problems relevant to valvular heart disease. Associated with each problem are problem factors and specifics used to determine a problem's severity. All the aspects in this patient attribute group are unquantified and should be left in patient description form, as if they had been related

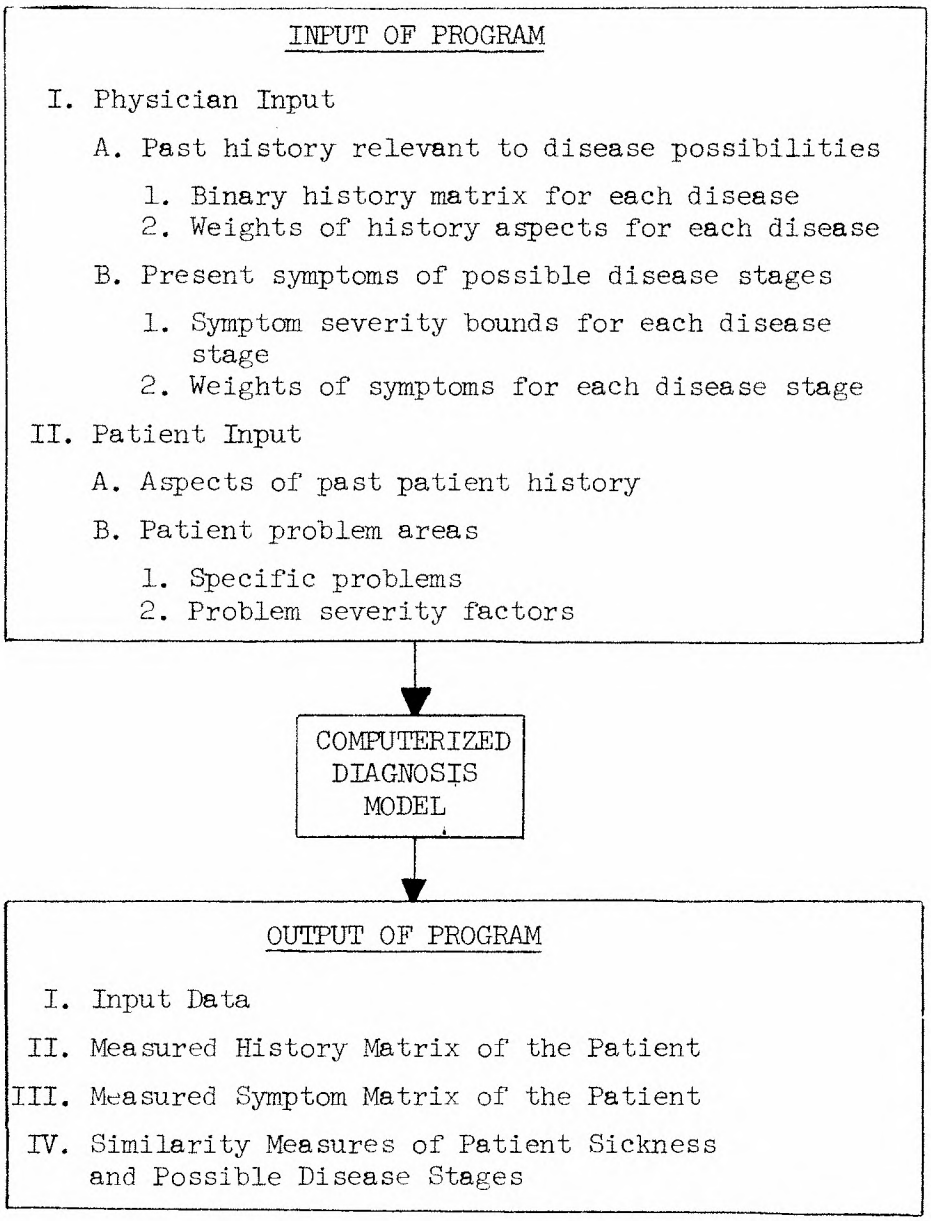

Figure 3. Diagram of Input and Output Information for the Computerized Diagnosis Model.

by the patient. For each patient's group of attributes or set of patient history and problems, the physician selects one or more possible disease stage candidates for the medical hypothesis. It should be noted that the physician's definition of symptoms must be consistent with those used for the computerized model.

A comparison between mock physician hypotheses and the results of the computerized decision model involving the previous medical example is presented in Table 18. One column of this table corresponds to the mock physician hypotheses, while another column corresponds to the hypotheses reached by the fuzzy decision model. The model and physician concurred on four of the hypothesis decisions, with the physician's hypothesis consistently being one of the six most likely disease stage candidates. This is illustrated in Table 18 in the column corresponding to the ranking of the physician's hypothesis. The last column of Table 18 corresponds to the rank of the physician hypothesis regardless of disease stage development. This would be similar to results of past modelling techniques where a disease's development is assumed constant.

Certainly the results in Table 18 illustrate discrepancies and problems with the proposed model. These discrepancies are probably due to a variety of factors. First of all, the diseases and disease stages are very similar to one another and lack numerous or severe discriminating factors. Furthermore, measurement or weighting inaccuracies are evident when the validation data is analyzed. Symptoms relevant to aortic and mitral valve diseases, such as angina, dyspnea, and syncope, are not specifically pertinent to tricuspid valve disease. Unfortunately, such symptoms are not negatively weighted enough for

Table 18. Comparison Between Physician and Computer Aided Hypotheses.

Patient Number	Physician Hypothesis-K	Computer Aided Hypothesis	Similarity Rank of K	Highest Rank for Any Stage-K
1	Severe Mitral Stenosis	Severe Mitral Stenosis	1	1
2	Intermediate Aortic Insufficiency	Intermediate Aortic Insufficiency	1	1
3	Severe Aortic Stenosis and/or Intermediate Mitral Stenosis	Intermediate Tricuspid Insufficiency	8/14	7/5
4	Severe Tricuspid Insufficiency and Intermediate Mitral Insufficiency	Severe Aortic Stenosis	4/10	3/7
5	Severe Mitral Stenosis and/or Severe Tricuspid Insufficiency	Severe Tricuspid Insufficiency	4/1	4/1
6	Intermediate Mitral Insufficiency	Mild Tricuspid Insufficiency	6	5
7	Severe Mitral Stenosis	Severe Mitral Stenosis	1	1
8	Severe Tricuspid Insufficiency	Intermediate Tricuspid Insufficiency	3	1
9	Intermediate Aortic Stenosis	Severe Aortic Insufficiency	6	6
10	Intermediate Aortic Insufficiency	Severe Aortic Insufficiency	2	1

tricuspid insufficiency and the resultant distance measure or similarity measure is somewhat erroneous. Symptoms differentiating aortic stenosis from aortic insufficiency must also receive measurement adjustments. Weights associated with aortic stenosis symptoms, such as angina and syncope, must be negatively increased for the stages of aortic insufficiency. Slight measurement revisions must also be made for the stages of tricuspid insufficiency and aortic insufficiency since the computerized hypothesis named the correct disease but the incorrect disease stage. Correction of these measurement discrepancies should result in a more realistic and accurate diagnosis model.

Since the initial problems confronted by this fuzzy modelling approach have been designated, additional modelling development should follow. Granted, as a diagnostic tool, the initial accuracy of this fuzzy diagnosis model is not extremely high, but the results are encouraging enough to warrant future model development along these lines. It must be pointed out that the decision model presented in this work did succeed in representing more of the influential aspects important to actual diagnosis decisions, than previous mathematical non-fuzzy models. Future research in model and data development via fuzzy systems theory should result in ever improving accuracy and preciseness.

CHAPTER VI

SUMMARY AND FUTURE RESEARCH

Research for this work comprises the initial phase of fuzzy model development in the area of medical decision making. Although fuzzy systems theory is discussed in prominent journals and references, applications to real life situations have been somewhat sparse and infrequent. This research has been presented in an attempt to alleviate this deficiency.

Medical decision making, strongly influenced by a variety of imprecise and difficult to quantify aspects, has presented itself as a fertile area for applications of fuzzy set theory. As illustrated in Chapter II, these aspects exist throughout the various decision processes of the physician. Whether in the diagnosis decision process, the test selection process, the treatment selection process, or the treatment of the patient, fuzziness underlies many of the decisions important to the well being of the patient. In the diagnosis decisions, fuzziness is found in the physician-patient interaction and in information concerning the state of the patient. This includes imprecise patient communication of symptoms, as well as abnormality or severity measurement of patient attributes. In the test selection process, many factors influencing the choosing of a test have been designated as vague or difficult to quantify. These involve the usefulness, costliness, need, convenience, and danger of the test. Selection of the proper

treatment is also constrained by such fuzzy variables, but their respective importance to the decision varies. Fuzziness also occurs in the actual treatment or treatment management of the patient, through instructions that can be vague, imprecise, or unquantified.

Accompanying the descriptions of these decision processes and their intrinsic fuzziness has been an evaluation of their past mathematical representations. Oversimplification and misrepresentation of fuzzy and nonfuzzy aspects has been a fault of past models. In an effort to correct past deficiences, fuzzy set theory and other conventional techniques have been suggested for more appropriate modelling.

The specific model developed in this research centers on the diagnosis decision process, which includes decisions upon medical hypothesis, preliminary diagnoses, and final diagnoses. Representations of patient information have been developed and quantification techniques for this information, including those of fuzzy set theory, have been offered. Organization and incorporation of patient data, both fuzzy and nonfuzzy, have been a major consideration of this modelling phase. Information including patient history, symptoms, signs, and test results can be input for the developed diagnosis decision model. The diagnosis decision model suggested involves the use of cluster analysis techniques to determine the most likely disease candidate for the diagnoses. Similarities between the patient's sickness and all possible disease stages should be calculated via this model to determine the most similar disease stage.

A medical example has been given in Chapter IV to illustrate an application of the developed diagnosis model. The information in the

example involves past history and symptoms of the patient and disease stages. Through the specifics of the example, a medical hypothesis of the patient in the disease area of valvular heart disease can be reached. Specific information concerning patient history, patient problems, problem factors, and fuzzy severity description have been given. Fuzzy sets have been created, their membership functions defined, and their role in symptom severity determination specified.

Using the data from the medical example, Chapter V and Appendix C have presented a computerized program of the diagnosis decision model. With this computerized version, validation of the fuzzy diagnosis model becomes feasible. As has been pointed out, patient files were not available for this model validation. Thus, an alternate technique, involving mock diagnoses by a physician, has been devised to determine the appropriateness of the proposed model. Comparisons between physician diagnoses and computer diagnoses with the same information have been made to determine this appropriateness. The results of this validation have shown initial promise for the developed model as well as the need for further model analysis.

Since this is an initial phase of fuzzy model development in the physician-patient decision area, future research offers many possibilities. These can be classified into three categories.

1. Designation and discussion of additional fuzzy aspects and considerations in the medical decision making field.
2. Standardization of medical definitions and influential factors relevant to physician decisions.
3. Improvements upon measurements of the existing model and

additional model development.

Each of these is in itself a formidable research task.

Although this research identified many of the major fuzzy aspects within the physician-patient interaction and relationship, further designation and interpretation of fuzzy considerations should be made. Refinements concerning the fuzzy aspects already designated, or additional factors of fuzziness warrant future study. In this research, factors influencing fuzziness within the patient-physician relations were outlined and discussed. Future work might center on influences outside this interactive relationship. Such research would include cultural influences upon fuzzy concepts, fuzzy interpretations, and assessment of fuzziness.

Possibly the most imposing research task involves attempts to standardize medical definitions and other medical factors. For future mathematical study in this field, more consistent definitions and interpretations of medical evidence must be developed. In this research, the possible symptoms for a given disease are fairly uniform from physician to physician and from reference to reference. However, the determination of severity levels of symptoms is based on less consistent information. Variations between physicians exist in the area of symptom definitions and determination of factors influencing symptom severity. One physician may have one criterion to determine the severity of a symptom, while another may use a completely different approach. This is the reason for assuming symptom definitions and severity factors listed in Appendix B. Uniform criteria must be developed in the future for determining symptom severity and other

diagnostic medical considerations. Such efforts are presently being undertaken by the American College of Physicians, but this will only be a first step in achieving the desired medical standardization.

The final category for possible future research involves refinement and further development of the fuzzy diagnosis model presented in this research. Refinements should be made on the proposed model to obtain more precise symptom and weight measurements. A broader base of opinion needs to be constructed for determination of these values. The Delphi Method for polling professional opinion might be a feasible approach to solve the problem. Once such expert opinion is obtained, further measurement refinement should be made. By validating the diagnosis model with more patient diagnoses, further information involving improper measurements should result. Sensitivity and statistical analysis might then be performed to evaluate this information. Since the example presented stops with the medical hypothesis, illustrating the decision process to the final diagnosis is desirable to verify other modelling suggestions involving sign designation and test result interpretation. Alternate physician decision processes, such as test and treatment selection, also need to be modelled via fuzzy set theory so that overall comprehensive diagnostic and treatment models might be feasible.

To accomplish this research, collaboration between the mathematical and medical fields was essential. Additional research, whether for educational or practical use, will demand even closer cooperation between individuals in their respective fields. Hopefully, this and other future mathematical models will be applied and used in the medical

profession. Initially, such computerized models may serve only as training tools for medical students and interns. As model accuracy and acceptance among physicians increase, medical decision models might eventually replace decision making aspects of physicians.

APPENDIX A

ELEMENTS OF FUZZY SET THEORY

Fuzzy Sets

As was stated earlier, a fuzzy set is defined as a set of elements in which there is no distinct boundary between the elements that belong to the set and those which do not belong. Hence, an element in a fuzzy set has a degree of membership rather than full membership or nonmembership.

Consider a set $X = \{x_i\}$ where x_i is a generic element of the set X. A fuzzy set A in X is a set of ordered pairs defined as

$$A = \{(x_i, u_A(x_i)) | x_i \in X\} . \tag{66}$$

The function $u_A(x_i)$ is called the membership function of x_i in A which maps X to the membership space M, where M = [0,1]. This is written symbolically as $\mu_A: X \rightarrow M$. If A is a nonfuzzy set, then $u_A(x_i) = 0$ or 1, according as $x_i \notin A$ or $x_i \in A$.

Example: Consider the clinical test involving white blood cell (leukocyte) count in a given patient. Let

$$X = \{x_i | x_i > 0\} . \tag{67}$$

Let A be a fuzzy set of highly abnormal white blood cell count (indication of infection). Define the membership function u_A, which maps X into [0,1] as follows:

$$u_A(x_i) = \begin{cases} 0 \ , & \text{for } x_i < 11{,}000 \\ (x_i - 11{,}000/9{,}000 \ , & \text{for } 11{,}000 \leq x_i \leq 20{,}000 \\ 1 \ , & \text{for } x_i > 20{,}000 \end{cases} \qquad (68)$$

As can be seen from this membership function, the abnormal range for leukocyte counts begins at 11,000.

Definitions Involving Fuzzy Sets

Empty Fuzzy Set

A fuzzy set is empty if and only if

$$u_A(x_i) = 0 \ , \quad \forall \ x_i \in X \ . \qquad (69)$$

Equality

Two fuzzy sets A and B are equal (A = B) if and only if

$$u_A(x_i) = u_B(x_i) \ , \quad \forall \ x_i \in X \ . \qquad (70)$$

Complement

The complement of a fuzzy set A, denoted by A, is defined by

$$u_{A'}(x_i) = 1 - u_A(x_i) \ , \quad \forall \ x_i \in X \ . \qquad (71)$$

Containment

Let A and B denote two fuzzy sets in X. A is contained in B or A is a subset of B($A \subset B$) if and only if

$$u_A(x_i) \leq u_B(x_i) \ , \quad \forall \ x_i \in X \ . \qquad (72)$$

Union

Let A be a fuzzy set in X with a membership function denoted by

$u_A(x_i)$, $\forall x_i \in X$. Let B be a fuzzy set in X with a membership function denoted by $u_B(x_i)$, $\forall x_i \in X$. The union of fuzzy sets A and B is a fuzzy set C, written as

$$C = A \cup B ,$$

where

$$u_C(x_i) = \max[u_A(x_i), u_B(x_i)] \quad , \forall x_i \in X \tag{73}$$

or, in abbreviated form

$$u_C = u_A \vee u_B . \tag{74}$$

Intersection

Let A and B be two fuzzy sets, denoted by the membership functions $u_A(x_i)$ and $u_B(x_i)$, $\forall x_i \in X$, respectively. The intersection of fuzzy sets A and B is a fuzzy set C, written as $C = A \cap B$, where

$$u_C(x_i) = \min [u_A(x_i), u_B(x_i)] , \forall x_i \in X \tag{75}$$

or, in abbreviated form

$$u_C = u_A \wedge u_B . \tag{76}$$

Convexity

A fuzzy set A is convex if and only if the sets Γ_α defined by

$$\Gamma_\alpha = \{x_i | u_A(x_i) \geq \alpha\} \tag{77}$$

are convex for all α in the interval (0,1]. Alternately, A is convex if and only if

$$u_A[\lambda x_1 + (1 - \lambda)x_2] \geq \min [u_A(x_1), u_A(x_2)] \tag{78}$$

for all x_1 and x_2 in X and all λ in [0,1].

Decision Making in a Fuzzy Environment

When making decisions in a fuzzy environment, the objectives of the decision, O, and the constraints of the decision, C, are fuzzy sets. The system under consideration, X, is nonfuzzy in nature, usually the real numbers. The objective functions and constraints are characterized by the membership functions $u_O(x_i)$ and $u_C(x_i)$ respectively. The membership function of the decision, $u_D(x_i)$ is

$$u_D(x_i) = u_O(x_i) \wedge u_C(x_i) \qquad \forall\, x_i \in X\ . \tag{79}$$

A decision can then be based on the membership function $u_D(x_i)$, $\forall x_i \in X$.

Example: The objective of this example is to find x_i such that it is substantially larger than 10. Thus,

$$\begin{aligned} u_O(x_i) &= 0 && , \quad \text{for } x_i < 10 \\ u_O(x_i) &= 1/(1 + (x_i - 10)^{-2}) && , \quad \text{for } x_i \geq 10\ . \end{aligned} \tag{80}$$

The constraint of the example is that x_i should be in the vicinity of 11. This is characterized by the membership function

$$u_C(x_i) = 1/(1 + (x_i - 11)^4) \quad , \quad \forall\, x_i \in X\ . \tag{81}$$

The membership function, $u_D(x_i) = u_O(x_i) \wedge u_C(x_i)$, $\forall\, x_i \in X$, where

$$u_D(x_i) = \begin{cases} \min[1/(1 + (x_i - 10)^{-2}),\ 1/1 + (x_i - 11)^4], & \text{for } x_i \geq 10 \\ 0 & \text{for } x_i < 10 \end{cases} \tag{82}$$

APPENDIX B

DESCRIPTION OF MEDICALLY DESIGNATED SYMPTOMS

Symptom and Symptom Severity Definitions

Exertional Dyspnea

Shortness of breath upon exertion; effort provokes inability to catch breath; problems breathing upon exertion.

Severity of Symptom - determined by minimum level of severity accompanying symptom.

Parosysmal Nocturnal Dyspnea - Orthopnea

Extreme shortness of breath while asleep; often awakes patient after few hours of retirement; rapidly relieved when patient sits up and breathes deeply; patient has need to get "fresh air"; can occur more than once a night; usually night after night.

Severity of Symptom - single level of severity assumed.

Cough*

For relief of tickling in throat; may occur upon exertion, recumbency or incessantly.

Severity of Symptom - severity of symptom depends upon activity of patient and productiveness of cough.

Hemoptsis*

Blood coughed up after tickling sensation in throat; blood tinged sputum; bloody sputum from respiratory tract.

---

*Note: Pulmonary Edema is assumed to be equivalent to severe levels of cough, accompanied with severe levels of hemoptsis.

Severity of Symptom - severity depends upon bloodiness of sputum.

Cyanosis

Purplish or greyish blue coloration of the skin and the mucous membranes due to deficient oxygenation of the blood.

Severity of Symptom - severity of the symptom depends upon the location of the symptom and degree of purpleness.

Malar Flush

Deep redness of cheeks and of lips, due to congestion of fine capillaries in the cheek bones.

Severity of Symptom - single level of severity assumed.

Jaundice

Yellowish or yellow-greenish discoloration of the skin and other body tissues and fluids by bile pigments.

Severity of Symptom - single level of severity assumed.

Weakness

Lack of energy, inability to move, lack of strength.

Severity of Symptom - severity determined by limitations on activity.

Fatigue

Exhaustion from exercise, laziness, tiredness from activity.

Severity of Symptom - severity determined by minimum activity causing fatigue.

Syncope

Symptom complex characterized by sudden transcient weakness, dizziness, and faintness.

Severity of Symptom - the severity of the symptom depends upon

the unawareness of the patient.

Palpitations

Consciousness of rapid, forceful, irregular beating of the heart, increased awareness of normal heart action, pounding heart sensations.

Severity of Symptom - severity of the symptom depends upon the minimum level of activity accompanying symptom.

Typical Anginal Pain

Angina pectoris is typically characterized by the occurrence of nonspecific substernal pain of varying discomfort. This pain usually radiates from the sternum upward to the neck and down the inner side of the lieft arm to the little or ring finger. The pain is usually brought on by exertion or intense mental emotion and lasts from a few seconds to a few minutes.

Severity of Symptom - the severity depends upon the minimum activity accompanying symptom and radiation of pain to other locations.

Epigastric Anginal Pain

Severe epigastric pain, occurring upon exertion, radiating back to the spine or upward to the sternum.

Severity of Symptom - single level of severity assumed.

Pain From Enlarged Liver

Pulsating non-radiating pain in upper abdomen, lower chest.

Severity of Symptom - single level of severity assumed.

Headaches

Pain associated in head and upper neck region.

Severity of Symptom - the severity depends upon the painfulness.

Peripheral Edema

Accumulation of excess fluids in extremities of body.

Severity of Symptom - the severity depends upon the location and continuity of the swelling.

Ascites (enlarged liver)

Slight bulging of the flanks, unnatural fullness of lower abdomen, accompanies right ventricular heart failure, preceded by swelling and edema in legs and abdomen wall.

Severity of Symptom - single level of severity assumed.

Weight Gain from Edema

Increase of weight, due to edema.

Severity of Symptom - severity of symptom dependent upon increase in weight.

Anorexia

Lack of desire to eat due to liver failure.

Severity of Symptom - single level of severity assumed.

Nausea and Vomiting

Stomach sickness, regurgitation, due to liver failure.

Severity of Symptom - single level of severity assumed.

APPENDIX C

COMPUTER PROGRAM

FOR THE DIAGNOSIS MODEL

Problem - Factor Sheet: Input Data

Patient Problem Areas

00. No more problems
01. Shortness of breath
02. Coughing
03. Blood in sputum
04. Purple skin color
05. Red face and lips
06. Yellow skin color
07. Weakness
08. Fatigue
09. Dizziness; Lightheadedness Confusion; Greying out; Unconsciousness
10. Irregular heartbeat
11. Pain
12. Swelling
13. Weight gain
14. Appetite loss
15. Nausea and vomiting

Location Factors

00. No location
01. Fingernails
02. Face
03. Lips
04. Ankles
05. Feet
06. Legs
07. Abdomen
08. Upper abdomen
09. Lower chest
10. Middle chest
11. Inner left arm
12. Arms
13. Back of neck
14. Head
15. Back

Continuity Factors

01. Continuous
02. Intermittent

Activities

00. No specific activity
01. Running
02. Climbing
03. Lifting heavy weight
04. Walking
05. Standing from sitting
06. Resting
07. Recumbency
08. Sleep

Symptom Severity Sheet: Input Data

Headache - Pain

1. Mild pain -	0.17
2. Pain -	0.50
3. Severe pain -	1.00

Hemoptsis - Bloodiness

1. A little blood -	0.23
2. Flects of blood -	0.39
3. Slight bleeding -	0.23
4. Large amounts of blood -	1.00
5. Profuse bleeding -	1.00
6. Blood -	0.54

Cough - Productivity

1. Productive -	1.00
2. Nonproductive -	0.00

Weight gain - Increasing weight

1. 5 lbs. -	0.28
2. 10 lbs. -	0.50
3. 15 lbs. -	0.75
4. 20 lbs. -	1.00
5. A little weight -	0.22
6. A lot of weight -	0.55

Cyanosis - Purpleness

1. Ashen -	0.09
2. Blue tinge -	0.32
3. Purple tinge -	0.51
4. Purple -	0.77
5. Very blue grey -	0.21
6. Very purple -	1.00

Syncope - Awareness of patient

1. Unconsciousness -	0.00
2. Confusion -	0.61
3. Greying out -	0.41
4. Dizziness -	0.78
5. Lightheadedness -	0.82

```
      PROGRAM MAIN(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT,TAPE2)
C
C
C  THIS PROGRAM INCORPORATES CLUSTER ANALYSIS AND FUZZY SET THEORY TO
C    REACH A DIAGNOSIS BASED ON PAST PATIENT HISTORY AND PATIENT
C   SYMPTOMS
C
      COMMON/ONE/V(8),VV(5,8),DV(5),XV(5,8)
      COMMON/TWO/X(20),XB(15,20),XL(15,20)
      COMMON/THREE/W(15,20),DX3M(15,20)
      COMMON/FOUR/XDIS(15),Y(15)
      COMMON/FIVE/YMIN(15),L(15)
      COMMON/SIX/B(1,15),IB(10,15)
      COMMON/SEVEN/FUNCT(7),IP(16),XMIN(15)
      COMMON/EIGHT/C(5,8)
C
C
C READ IN PHYSICIAN INPUT
C
C  READ IN WEIGHTS OF HISTORY
      READ(2,30)((C(I,J),J=1,8),I=1,5)
 30   FORMAT(8F3.0)
C
C  READ IN HISTORY MATRIX FOR EACH DISEASE
C
      READ(2,35)((VV(I,J),J=1,8),I=1,5)
 35   FORMAT(8F2.0)
C
C  READ IN UPPER SEVERITY AND LOWER SEVERITY BOUNDS FOR SYMPTOMS
C
      READ(2,37)((XL(I,J),I=1,15),J=1,20)
      READ(2,37)((XU(I,J),I=1,15),J=1,20)
 37   FORMAT(15F3.0)
C
C  READ IN SYMPTOM WEIGHTS
C
      READ(2,40)((W(I,J),J=1,20),I=1,15)
 40   FORMAT(20F3.0)
C
C  PRINT OUT INPUT IF DESIRED
C
      WRITE(6,90)
 90   FORMAT(3X,41HDO YOU WANT INFORMATION MATRICES PRINTED /24HENTER 1
     1(YES) OR 0 (NO).)
      READ(5,95)INF
 95   FORMAT(I1)
      IF(INF.NE.1)GO TO 130
C
C     PRINT OUT INPUT FOR DISEASES
C
      WRITE(6,100)((C(I,J),J=1,8),I=1,5)
 100  FORMAT(8F4.1)
      WRITE(6,105)((VV(I,J),J=1,8),I=1,5)
 105  FORMAT(8F4.1)
      WRITE(6,110)((XU(I,J),J=1,20),I=1,15)
      WRITE(6,110)((XL(I,J),J=1,20),I=1,15)
      WRITE(6,110)((W(I,J),J=1,20),I=1,15)
 110  FORMAT(20F4.1)
 130  CONTINUE
      WRITE(6,137)
 137  FORMAT(///)
C
C
C
C  READ IN THE NUMBER OF DISEASES TO BE DIAGNOSED
C
C
      WRITE(6,135)
 135  FORMAT(45HHOW MANY PATIENT DISEASES WILL BE DIAGNOSED  /17HREAD IN
     1 I2 FORMAT)
      READ(5,140)NUM
 140  FORMAT(I2)
      DO 999 N=1,NUM
C
C INITIALIZE PATIENT HISTORY MATRIX
      DO 150 J=1,8
      V(J)=0.0
 150  CONTINUE
C
```

```
      BIGNUM=1.9
C
C INITIALIZE PATIENT SYMPTOM MATRIX
      DO 160 J=1,20
      X(J)=BIGNUM
 160  CONTINUE
C
C  INITIALIZE OTHER VARIABLES
      DO 180 I=1,15
      DO 170 J=1,20
      QYSM(I,J)=0.0
 170  CONTINUE
      Y(I)=0.0
 180  CONTINUE
      DO 185 I=1,10
      DO 185 K=1,15
      IB(I,K)=0
      XMIN(K)=0.0
 185  CONTINUE
C
C
C  READ IN PAST MEDICAL HISTORY OF PATIENT
      WRITE(6,190)
 190  FORMAT(///)
C
      WRITE(6,210)
 210  FORMAT(10X,21HPATIENTS PAST HISTORY//3X,50HREAD IN PATEINTS PAST
     IDISEASES USING APPROPRIATE NUMBER//10X,17H1-RHEUMATIC FEVER//10X,1
     20H2-SYPHILIS//10X,30H3-RHEUMATIC FEVER AND SYPHILIS//10X,24H4-NONE
     3 OF THESE DISEASES//)
C
C
C  READ IN PAST DISEASES
      READ(5,215)IDS
 215  FORMAT(I1)
      IF(IDS.NE.1)GO TO 220
      V(1)=1.0
      GO TO 240
 220  IF(IDS.NE.2)GO TO 225
      V(2)=1.0
      GO TO 240
 225  IF(IDS.NE.3)GO TO 230
      V(1)=1.0
      V(2)=1.0
      GO TO 240
 230  V(3)=1.0
 240  CONTINUE
      IF(V(3).NE.1.0)GO TO 245
C
C  DETERMINE IF SYMPTOMS OF UNDIAGNOSED DISEASES HAVE EXISTED
C
      WRITE(6,242)
 242  FORMAT(///31HHAS THE PATIENT EVER HAD EITHER//3X,37HOF THE FOLLOWI
     1NG GROUPS OF SYMPTOMS  //3X,40HIF SO DESIGNATE WITH APPROPRIATE NU
     2MBERS//10X,28H1-HIGH FEVER AND SORE JOINTS//10X,26H2-PINK SKIN RAS
     3H AND SORES//10X,18H3-NEITHER OF THESE//)
      READ(5,243)ISTM
 243  FORMAT(I1)
      IF(ISTM.NE.1)GO TO 244
      V(1)=1.0
      V(3)=0.0
      GO TO 246
 244  IF(ISTM.NE.2)GO TO 246
      V(2)=1.0
      V(3)=0.0
 246  CONTINUE
 245  CONTINUE
C
C READ IN PATIENTS SEX
C
      WRITE(6,250)
 250  FORMAT(////45HREAD IN PATIENTS SEX USING APPROPRIATE NUMBER//10X,6
     1H1-MALE//10X,8H2-FEMALE)
      READ(5,255)ISEX
 255  FORMAT(I1)
      IF(ISEX.NE.1)GO TO 260
      V(4)=1.0
      GO TO 265
 260  V(5)=1.0
 265  CONTINUE
```

```
C
C  READ IN PATIENTS AGE GROUP
C
      WRITE(6,269)
  269 FORMAT(///53HREAD IN PATIENT AGE CATEGORY USING APPROPRIATE NUMBER
     1//10X,12H1-(UNDER 30)//10X,9H2-(30-60)//10X,11H3-(OVER 60))
      READ(5,275)IGE
  275 FORMAT(I1)
      IF(IGE.NE.1)GO TO 280
      V(6)=1.0
      GO TO 300
  280 IF(IGE.NE.2)GO TO 290
      V(7)=1.0
      GO TO 300
  290 V(8)=1.0
  300 CONTINUE
C
C   PATIENTS HISTORY MATRIX IS NOW KNOWN
C
C  DETERMINE SIMILARITY BETWEEN PATIENTS HISTORY AND DISEASE I
C
      DO 330 I=1,5
      DV(I)=0.0
      DO 320 J=1,8
      XV(I,J)=(C(I,J)*(V(J)-VV(I,J)))**2
      DV(I)=DV(I)+XV(I,J)
  320 CONTINUE
      DV(I)=DV(I)**.5
  330 CONTINUE
C
C
      K=0
      MM=1
C
C
C READ IN THE PATIENT PROBLEMS
C
  500 CONTINUE
      WRITE(6,503)
  503 FORMAT(////)
      WRITE(6,505)
  505 FORMAT(52HREAD IN A PATIENT PROBLEM AREA USING THE APPROPRIATE/2X,
     138HNUMBER FROM THE PROBLEM FACTOR SHEET  /13HUSE I2 FORMAT/)
      READ(5,510)IP(MM)
  510 FORMAT(I2)
      IF(IP(MM).EQ.0)GO TO 600
      K=IP(MM)
      IB(1,K)=1
C
C  READ IN LOCATIONS OF PROBLEM
C
      WRITE(6,515)
  515 FORMAT(2X,41HHOW MANY LOCATIONS DOES THIS PROBLEM HAVE/)
      READ(5,520)NLOC
  520 FORMAT(I1)
      IF(NLOC.EQ.0)GO TO 530
      WRITE(6,522)
  522 FORMAT(3X,25HREAD IN PROBLEM LOCATIONS/3X,31HDESIGNATE ONE LOCATIO
     1N PER LINE/3X,13HUSE I2 FORMAT/)
      MLOC=NLOC+5
      READ(5,523)(IB(I,K),I=6,MLOC)
  523 FORMAT(I2)
C
C DETERMINE IF PROBLEM IS CONTINUOUS OR INTERMITTENT
  530 CONTINUE
      WRITE(6,532)
  532 FORMAT(3X,29HREAD IN CONTINUITY OF PROBLEM/3X,39HUSING NUMBER FROM
     1 PROBLEM- FACTOR SHEET/3X,13HUSE I2 FORMAT)
      READ(5,535)IB(3,K)
  535 FORMAT(I2)
C DETERMINE ACTIVITY ACCOMPANYING INTERMITTENT PROBLEM OCCURRENCE
      IF(IB(3,K).EQ.1)GO TO 550
      WRITE(6,537)
  537 FORMAT(3X,51HREAD IN MINIMUM ACTIVITY OR ACTIVITIES ACCOMPANYING/3
     1X,18HSYMPTOM OCCURRENCE//3X,33HUSE APPROPRIATE NUMBER FROM SHEET/3
     2X,47HENTER TWO ACTIVITY NUMBERS,ONE NUMBER PER LINE//3X,14HUSE I2
     3 FORMAT)
      READ(5,540)(IB(J,K),J=4,5)
  540 FORMAT(I2)
```

```
 558  CONTINUE
C READ IN SEVERITY SPECIFIC VALUES FOR CERTAIN PROBLEMS
      WRITE(6,552)
 552  FORMAT(3X,47HENTER SEVERITY VALUE FOR EACH REQUESTED SYMPTOM/3X,26
     1HUSE SYMPTOM SEVERITY SHEET/3X,17HUSE F(4.2) FORMAT//)
      IF(K.NE.2)GO TO 555
C  DETERMINE SEVERITY FOR COUGH
      WRITE(6,553)
 553  FORMAT(3X,5HCOUGH//)
      READ(5,554)B(2,2)
 554  FORMAT(F4.2)
      GO TO 595
 555  CONTINUE
      IF(K.NE.3)GO TO 560
C DETERMINE SEVERITY FOR HEMOPTSIS
      WRITE(6,556)
 556  FORMAT(3X,9HHEMOPTSIS//)
      READ(5,558)B(2,3)
 558  FORMAT(F4.2)
      GO TO 595
 560  CONTINUE
C FIND SEVERITY OF CYANOSIS
      IF(K.NE.4)GO TO 570
      WRITE(6,562)
 562  FORMAT(3X,8HCYANOSIS/)
      READ(5,565)B(2,4)
 565  FORMAT(F4.2)
      GO TO 595
 570  CONTINUE
C     FIND SEVERITY OF SYNCOPE
      IF(K.NE.9)GO TO 575
      WRITE(6,572)
 572  FORMAT(3X,7HSYNCOPE/)
      READ(5,574)B(2,9)
 574  FORMAT(F4.2)
      GO TO 595
 575  CONTINUE
C      FIND  SEVERITY OF HEADACHE
      IF(K.NE.11)GO TO 585
      DO 578 I=6,10
      IF(IB(I,K).EQ.13.OR.IB(I,K).EQ.14)GO TO 580
  578 CONTINUE
      GO TO 585
 580  CONTINUE
C FIND SEVERITY OF HEADACHE
      WRITE(6,581)
 581  FORMAT(3X,8HHEADACHE)
      READ(5,583)B(2,11)
 583  FORMAT(F4.2)

      GO TO 595

 585  CONTINUE
C  FIND SEVERITY OF WEIGHT GAIN
 590  IF(K.NE.13)GO TO 591
      WRITE(6,592)
 592  FORMAT(3X,11HWEIGHT GAIN)
      READ(5,594)B(2,13)
 594  FORMAT(F4.2)
      GO TO 595
 591  WRITE(6,597)
 597  FORMAT(//3X,24HNO SEVERITY LEVEL NEEDED//)
 595  CONTINUE
      MM=MM+1
      GO TO 500
 600  CONTINUE
      K=0
C DETERMINE MEMBERSHIP FUNCTIONS FOR FUZZY SET DENOTING STRENUOUSNESS OF
C   ACTIVITY
      FUNCT(1)=1.0
      FUNCT(2)=.93
      FUNCT(3)=.96
      FUNCT(4)=.36
      FUNCT(5)=.14
      FUNCT(6)=.12
      FUNCT(7)=0.0
C
C
C
C  DETERMINE MINIMUM LEVEL OF ACTIVITY ACCOMPANYING SYMPTOM
```

```
      DO 615 K=1,15
      IF(IB(4,K).EQ.0.AND.IB(5,K).EQ.0)GO TO 615
      IF(IB(4,K).EQ.0.OR.IB(5,K).EQ.0)GO TO 605
      IF(IB(4,K).EQ.8.OR.IB(5,K).EQ.8)GO TO 605
      A1=IB(4,K)
      A2=IB(5,K)
      IF(FUNCT(A1).LT.FUNCT(A2))GO TO 601
      XMIN(K)=FUNCT(A2)
      GO TO 603
 601  XMIN(K)=FUNCT(A1)
 603  CONTINUE
      GO TO 615
 605  IF(IB(4,K).EQ.0)GO TO 607
      IF(IB(4,K).EQ.8)GO TO 607
      A3=IB(4,K)
      XMIN(K)=FUNCT(A3)
      GO TO 615
 607  IF(IB(5,K).EQ.8)GO TO 615
      A3=IB(5,K)
      XMIN(K)=FUNCT(A3)
 615  CONTINUE
C
C
C
C DETERMINE PATIENT SYMPTOM MATRIX
C
C    DETERMINE IF DYSPNEA IS PRESENT
      IF(IB(1,1).NE.1)GO TO 620
      IF(IB(4,1).EQ.8.OR.IB(5,1).EQ.8) GO TO 616
C  EXERTIONAL DYSPNEA IS PRESENT
      X(1)=1.0-XMIN(1)
      GO TO 618
 616  X(2)=1.0
C  PAROXYSMAL NOCTURNAL DYSPNEA IS PRESENT
      IF(IB(4,1).EQ.0.OR.IB(5,1).EQ.0)GO TO 620
      X(1)=1.0-XMIN(1)
 618  CONTINUE
 620  IF(IB(1,2).NE.1)GO TO 625
      X(3)=B(2,2)
C  COUGH IS PRESENT
 625  IF(IB(1,3).NE.1)GO TO 630
      IF(IB(3,3).EQ.1)GO TO 629
      IF(B(2,2).EQ.1.0)GO TO 627
      IF(IB(4,3).GT.5.OR.IB(5,3).GT.5)GO TO 626
      X(4)=.14
      GO TO 630
 626  X(4)=.62
      GO TO 630
 627  IF(IB(4,3).GT.5.OR.IB(5,3).GT.5)GO TO 628
      X(4)=.7
      GO TO 630
 628  X(4)=.89
      GO TO 630
 629  X(4)=1.0
C  HEMOPTSIS IS PRESENT
 630  IF(IB(1,4).NE.1)GO TO 635
C  CYANOSIS IS PRESENT
      DO 633 I=6,10
      IF(IB(I,4).NE.1.OR.IB(I,4).NE.2)GO TO 631
      X(5)=B(2,4)/2.
      GO TO 633
 631  X(5)=(B(2,4)/2.)+.5
      GO TO 635
 633  CONTINUE
 635  IF(IB(1,5).NE.1) GO TO 640
      X(6)=1.0
C  MALAR FLUSH IS PRESENT
 640  IF(IB(1,6).NE.1)GO TO 645
C  JAUNDICE IS PRESENT
      X(7)=1.0
 645  IF(IB(1,7).NE.1)GO TO 650
      X(8)=1.0-XMIN(7)
C  WEAKNESS IS PRESENT
 650  IF(IB(1,8).NE.1)GO TO 655
C FATIGUE IS PRESENT
      X(9)=1.0-XMIN(8)
 655  IF(IB(1,9).NE.1)GO TO 660
      X(10)=B(2,9)
C  SYNCOPE IS PRESENT
```

```
  660  IF(IB(I,10).NE.1)GO TO 665
       X(11)=1.0-XMIN(10)
  665  IF(IB(I,11).NE.1)GO TO 690
       DO 685 J=6,10
       IF(IB(I,11).EQ.10)GO TO 666
       IF(IB(I,11).EQ.7.OR.IB(I,11).EQ.15)GO TO 670
       IF(IB(I,11).EQ.8.OR.IB(I,11).EQ.9)GO TO 675
       IF(IB(I,11).EQ.13.OR.IB(I,11).EQ.14)GO TO 680
       GO TO 685
C  TYPICAL ANGINA IS PRESENT
  666  DO 667 J=6,10
       IF(IB(J,11).EQ.11)GO TO 668
       X(12)=(1.0-XMIN(11))/2.
       GO TO 667
  668  X(12)=(1.0-XMIN(11))/2.+.5
       GO TO 685
  667  CONTINUE
C  EPIGASTRIC ANGINAL PAIN IS PRESENT
  670  IF(XMIN(11).LT..1)GO TO 685
       X(13)=1.0
       GO TO 685
C PAIN DUE TO ENLARGED LIVER IS PRESENT
  675  X(14)=1.0
       GO TO 685
C  HEADACHES ARE PRESENT
  680  X(15)=B(2,11)
  685  CONTINUE
  690  IF(IB(1,12).NE.1)GO TO 700
       DO 694 M=6,10
       IF(IB(M,12).EQ.7)GO TO 698
  694  CONTINUE
C PERIPHERAL EDEMA IS PRESENT
  695  DO 696 MN=6,10
       IF(IB(MN,12).NE.4.OR.IB(MN,12).NE.5.OR.IB(MN,12).NE.8)GO TO 697
  696  CONTINUE
       IF(B(3,12).EQ.1.0)GO TO 693
       X(16)=.24
       GO TO 700
  693  X(16)=.41
       GO TO 700
  697  IF(B(3,12).EQ.1.0)GO TO 699
       X(16)=.79
       GO TO 700
  699  X(16)=1.0
       GO TO 700
C ASCITES IS PRESENT
  698  X(17)=1.0
       GO TO 695
  700  CONTINUE
       IF(IB(1,13).NE.1)GO TO 710
       X(18)=B(2,13)
C    WEIGHT GAIN IS PRESENT
  710  IF(IB(1,14).NE.1)GO TO 720
       X(19)=1.0
  720  IF(IB(1,15).NE.1)GO TO 730
       X(20)=1.0
  730  CONTINUE
C
C THE SIMILARITY BETWEEN THE PATIENTS SYMPTOMS AND EACH DISEASE SYMPTOMS
C IS DETERMINED
       DO 850 I=1,15
       DO 840 J=1,20
C
C DO NOT CONSIDER SYMPTOMS WHICH HAVE NOT BEEN DESIGNATED AS PRESENT OR
C ABSENT
C
       IF(X(J).EQ.BIGNUM)GO TO 840
C
       IF(X(J).GT.XU(I,J))GO TO 810
       IF(X(J).LT.XL(I,J))GO TO 820
C
C  THE PATIENT SYMPTOM IS IN THE PROPER SEVERITY INTERVAL FOR DISEASE
       DYSM(I,J)=0.0
C
       GO TO 830
C
C PATIENTS SYMPTOM IS ABOVE LEVELS FOR DISEASE STAGE
C
  810  DYSM(I,J)=(W(I,J)*(X(J)-XU(I,J)))**2
       GO TO 830
```

```
C PATIENTS SYMPTOM IS BELOW LEVELS FOR DISEASE STAGE
C
  820   DYSM(I,J)=(W(I,J)*(XL(I,J)-X(J)))**2
  830   CONTINUE
        XDIS(I)=XDIS(I)+DYSM(I,J)
  840   CONTINUE
C
        XDIS(I)=XDIS(I)**.5
  850   CONTINUE
C
C
C SUM THE SIMILARITIES BETWEEN PATIENT HISTORY AND PATIENT SYMPTOMS
C  FOR EACH DISEASE
C
        J=0
        DO 860 K=1,15,3
        J=J+1
        Y(K)=DV(J)+XDIS(K)
        Y(K+1)=DV(J)+XDIS(K+1)
        Y(K+2)=DV(J)+XDIS(K+2)
  860   CONTINUE
C
        DO 735 J=1,20
        IF(X(J).NE.BIGNUM)GO TO 735
        X(J)=0.0
  735   CONTINUE
        WRITE(6,740)
  740   FORMAT(///14HHISTORY NUMBER,10X,19HPRESENCE OF HISTORY)
        DO 745 KK=1,8
        WRITE(6,742)KK,V(KK)
  742   FORMAT(5X,I2,20X,F3.0)
  745   CONTINUE
        WRITE(6,750)
  750   FORMAT(///14HSYMTPOM NUMBER,10X,17HLEVEL OF SEVERITY)
        DO 760 II=1,20
        WRITE(6,755)II,X(II)
  755   FORMAT(5X,I2,20X,F6.3)
  760   CONTINUE
C ORDER THE SIMILARITIES BETWEEN THE PATIENTS SICKNESS AND DISEASES
C
C TO ATTAIN THIS , FIND THE MOST DISSIMILAR DISEASE STAGE
C
        YMAX=Y(1)
        KMAX=1
  870   DO 880 K=2,15
        IF(YMAX.GE.Y(K))GO TO 880
        YMAX=Y(K)
        KMAX=K
  880   CONTINUE
C
C FIND THE MOST SIMILAR DISEASE AND ORDER THE DISEASES WITH RESPECT
C  TO SIMILARITY WITH THE PATIENTS SICKNESS
C
        DO 900 J=1,15
        L(J)=KMAX
        YMIN(J)=YMAX
        DO 895 I=1,15
C
C  ELIMINATE DISEASES ALREADY ORDERED
C
        IF(J.EQ.1)GO TO 891
        NN=J-1
        DO 890 M=1,NN
        IF(L(M).EQ.I)GO TO 895
  890   CONTINUE
  891   CONTINUE
        IF(YMIN(J).LT.Y(I))GO TO 895
        YMIN(J)=Y(I)
        L(J)=I
  895   CONTINUE
  900   CONTINUE
```

```
C
C  PRINT A TABLE OF THE ORDERED SIMILARITIES
C
      WRITE(6,910)
 910  FORMAT(//,52HTHE SIMILARITY MEASURES BETWEEN THE PATIENTS DISEASE/
     146HAND DISEASE I ARE GIVEN IN THE FOLLOWING TABLE///20X,21HTABLE O
     2F SIMILARITIES//)
      WRITE(6,920)
 920  FORMAT(10X,8HORDER OF,10X,7HDISEASE,10X,10HSIMILARITY/10X,10HSIMIL.
     1ARITY,8X,6HNUMBER,11X,7HMEASURE/)
      DO 930 J=1,15
      WRITE(6,925)J,L(J),YMIN(J)
 925  FORMAT(14X,I2,16X,I2,14X,F8.5)
 930  CONTINUE
 999  CONTINUE
      STOP
      END
```

BIBLIOGRAPHY

References - Medical

1. Bernstein, L., Bernstein, R. S., and Dana, R. H., Interviewing: A Guide for Health Professionals, Prentice-Hall, Inc., New York, 1974.

2. Conn, H. L., Jr., and Horwitz, O., Cardiac and Vascular Diseases, Volume I, Lea and Febiger, Philadelphia, 1971.

3. Feinstein, A. R., Clinical Judgment, Robert E. Krieger Co., Huntington, New York, 1967.

4. Feinstein, A. R., "An Analysis of Diagnostic Reasoning: The Domains and Disorders of Clinical Macrobiology", Yale Journal of Biology and Medicine 46, 212-232, 1973.

5. Feinstein, A. R., "An Analysis of Diagnostic Reasoning: The Strategy of Intermediate Decisions", Yale Journal of Biology and Medicine 46, 264-283, 1973.

6. Frenay, A. C., Understanding Medical Terminology, The Catholic Hospital Association, St. Louis, 1973.

7. Herrmann, G. R., Diseases of the Heart and Arteries, C. V. Mosby Co., St. Louis, 1952.

8. Hurst, J. W., ed., The Heart, Arteries and Veins, McGraw-Hill, New York, 1974.

9. Judge, R. D., ed., Physical Diagnosis: A Physiologic Approach to the Clinical Examination, Little, Brown, and Company, Boston, 1963.

10. Krieg, A. F., Gambino, R., and Galen, R. S., "Why Are Clinical Laboratory Tests Performed? When are They Valid?", Journal of American Medical Association 233(1), 76-78, 1975.

11. Krupp, M. A., and Chatton, M. J., Current Medical Diagnosis and Treatment, Lange Medical Publications, Los Altos, California, 1973.

12. Oram, S., Clinical Heart Disease, F. A. Davis Co., Philadelphia, 1971.

13. Reichek, N., Shelburne, J. C., and Perloff, J. K., "Clinical Aspects of Rheumatic Valvular Disease", Progress in Cardiovascular Diseases 15: 5, 491-533, 1973.

14. Scherf, D., and Boyd, L. J., Cardiovascular Diseases, Grune and Stratton, New York, 1958.

15. Stevenson, I., The Diagnostic Interview, Harper and Row, New York, 1971.

References - Modelling

16. Anderberg, M. R., Cluster Analysis for Applications, Academic Press, New York, 1973.

17. Barnoon, S., and Wolfe, H., Measuring the Effectiveness of Medical Decisions: An Operations Research Approach, Charles C. Thomas, Springfield, Illinois, 1972.

18. Bellman, R., "Mathematical Models of the Mind", Mathematical Biosciences 1, 287-304, 1967.

19. Bellman, R., and Smith, C. P., Simulation in Human Systems: Decision-Making in Psychotherapy, John Wiley and Sons, New York, 1973.

20. Betaque, N. E., and Gorry, A. G., "Automated Judgmental Decision Making for a Serious Medical Problem", Management Science 17B, 421-434, 1971.

21. Bouckaert, A., and Beckers, C., "Computer-Aided Differential Diagnosis of Goitres in Internal Medicine", Biochemical Computing 4, 1-17, 1973.

22. Chesler, L. G., Hershdorfer, A. M., and Lincoln, T. L., "The Use of Information in Clinical Problem Solving: A Framework for Analysis", Mathematical Biosciences 8, 83-108, 1970.

23. Croft, D. J., "Is Computerized Diagnosis Possible?", Computers and Biomedical Research 5, 351-367, 1972.

24. Cumerbatch, J., and Heaps, H. S., "Application of a Non-Bayesian Approach to Computer-Aided Diagnosis for Upper Abdominal Pain", Biomedical Computing 4, 105-115, 1973.

25. Cumerbatch, J., Leung, K. V., and Heaps, H. S., "A Non-Probablistic Method for Automated Medical Diagnosis", International Journal of Biochemical Computing 5, 133-146, 1974.

26. Duran, B. S., and Lesis, T. O., "An Application of Cluster Analysis to the Construction of a Diagnostic Classification", Computers in Biology and Medicine 4, 183-188, 1974.

27. Esogbue, A. O., Aggarwal, J., and Kajulgi, D., "Computer Aided Anesthesia Administration", International Journal of Biomedical Computing 7, 271-288, 1976.

28. Freemon, F. R., "Medical Diagnosis: Comparison of Human Computer Logic", Biomedical Computing 3, 217-221, 1972.

29. Gleser, M. A., and Collen, M. F., "Towards Automated Medical Decisions", Computers and Biomedical Research 5, 180-189, 1972.

30. Gorry, G. A., "Strategies for Computer Aided Diagnosis", Mathematical Biosciences 2, 293-318, 1968.

31. Gorry, G. A., and Barnett, G. O., "Sequential Diagnosis by Computer", Journal of American Medical Association 205, 849-854, 1968.

32. Gorry, G. A., "Computer-Assisted Clinical Decision Making", Methods of Information in Medicine 12, 45-53, 1973.

33. Henschke, U. K., and Flehinger, B. J., "Decision Theory in Cancer Therapy", Cancer 20:11, 1819-1826, 1967.

34. Jacquez, J. A., ed., Computer Diagnosis and Diagnostic Methods, Charles C. Thomas, Springfield, Illinois, 1972.

35. Knapp, D. E., and Oeltjen, P. D., "Benefits-to-Risks Ratio in Physician Drug Selection", American Journal of Public Health 62, 1346-1347, 1972.

36. Ledley, R. S., and Lusted, L. B., "Reasoning Foundations of Medical Diagnosis", Science 130: 3366, 9-21, 1959.

37. Leroy, A., "The Use of Information in Diagnostic Decisions", Acta Psychiatrica Scandinavica 49, 405-417, 1973.

38. Lively, W. M., Szygenda, S. A., and Mize, C. E., "Modelling Techniques for Medical Diagnosis I. Heuristics and Learning Programs in Selected Neonatal Hepatic Disease", Computers and Biomedical Research 6, 393-410, 1973.

39. Lusted, L. B., Introduction to Medical Decision Making, Thomas, Springfield, Illinois, 1968.

40. Mellergard, M., and Leroy, A., "Diagnostic Uncertainty and Therapeutic Decision", Acta Psychiatrica Scandinavica 49, 570-579, 1973.

41. Muic, V., Petres, J. J., and Telisman, Z., "Validity of a Diagnostic Test by a Single Function", Methods of Information in Medicine 12, 244-248, 1973.

42. Norusis, M. J., and Jacquez, J. A., "Diagnosis. I. Symptom Nonindependence in Mathematical Models for Diagnosis", Computers and Biomedical Research 8, 156-172, 1975.

43. Norusis, M. J., and Jacquez, J. A., "Diagnosis. II. Diagnostic Models Based on Attribute Clusters: A Proposal and Comparisons", Computers and Biochemical Research 8, 173-188, 1975.

44. Parker, S. G., and Kassirer, J. P., "Therapeutic Decision Making: A Cost Benefit Analysis", New England Journal of Medicine 293, 229-234, 1975.

45. Rector, A. L., and Ackerman, E., "Rules for Sequential Diagnosis", Computers and Biomedical Research 8, 143-155, 1975.

46. Schwartz, W. B., Gorry, G. A., Kassirer, J. P., and Essig, A., "Decision Analysis and Clinical Judgment", The American Journal of Medicine 55, 459-472, 1973.

47. Sheldon, A., Baker, F., and McLaughlin, C. P., eds., Systems and Medical Care, MIT Press, Cambridge, Mass., 1970.

48. Vanderplas, J. M., "A Method for Determining Probabilities for Correct Use of Bayes's Theorem in Medical Diagnosis", Computers and Biomedical Research 1, 215-220, 1967.

49. Warner, H. R., Toronto, A. F., Veasey, L. G., and Stephenson, R., "A Mathematical Approach to Medical Diagnosis", Journal of American Medical Association, 177: 3, 177-184, 1961.

50. White, C. C., "A New Physician-Patient Interaction Model for Medical Diagnosis and Treatment", Proceedings of MSAC-74, Milwaukee, Wisconsin, 1974.

References - Fuzzy Theory

51. Bellman, R., Kalaba, R., and Zadeh, L. A., "Abstraction and Pattern Classification", Journal of Mathematical Analysis and Applications 13, 1-7, 1966.

52. Bellman, R., and Zadeh, L. A., "Decision Making in a Fuzzy Environment", Management Science 17: 4, B141-B164, 1970.

53. Chang, S. L., "On Risk and Decision Making in a Fuzzy Environment", Fuzzy Sets and their Applications to Cognitive and Decision Processes, Zadeh, L. A., Fu, K., Tanaka, K., and Shimura, M., eds., Academic Press, Inc., New York, 1975.

54. DeLuca, A., and Termini, S., "A Definition of a Nonprobabilistic Entropy in the Setting of Fuzzy Sets Theory", Information and Control 20, 301-312, 1972.

55. Esogbue, A. O., Dynamic Programming and Fuzzy Allocation Processes, Technical Memorandum No. 202, Operations Research Department, Case Western University, 1970.

56. Esogbue, A. O., "On the Application of Fuzzy Allocation Theory to the Modelling of Cancer Research Appropriation Process", Modern Trends in Cybernetics and Systems, Rose, J. and Belciu, C., eds., Abacus Press, London (in press).

57. Goguen, J. A., "L-Fuzzy Sets", Journal of Mathematical Analysis and Applications 18, 145-174, 1967.

58. Lochen, M., and Badre, A. N., "On the Precision of Adjectives which Denote Fuzzy Sets", Journal of Cybernetics 4: 1, 49-60, 1974.

59. Kochen, M., "Application of Fuzzy Sets in Psychology", Fuzzy Sets and Their Applications to Cognitive and Decision Processes, Zadeh, L. A., Fu, K., Tanaka, K., and Shimura, M., eds., Academic Press, Inc., New York, 1975.

60. Zadeh, L. A., "Fuzzy Sets", Information and Control 8, 338-353, 1965.

61. Zadeh, L. A., "Shadows of Fuzzy Sets", Problems in Information Transmission 2: 1, 27-34, 1966.

62. Zadeh, L. A., "Fuzzy Algorithms", Information and Control 12, 94-102, 1968.

63. Zadeh, L. A., "Quantitative Fuzzy Semantics", Information Sciences 3, 159-176, 1971.

64. Zadeh, L. A., "Similarity Relations and Fuzzy Orderings", Information Sciences 3, 177-200, 1971.

65. Zadeh, L. A., Fu, K., Tanaka, K., and Shimura, M., eds., Fuzzy Sets and Their Applications to Cognitive and Decision Processes, Academic Press, Inc., New York, 1975.